DOCUMENT RESUME

ED 226 027 TM 830 055

AUTHOR	Wainer, Howard
TITLE	Multivariate Display for Quipus to Faces. Program Statistics Research.
INSTITUTION	Bureau of Social Science Research, Inc., Washington, D.C.; Educational Testing Service, Princeton, N.J.
SPONS AGENCY	National Science Foundation, Washington, D.C.
REPORT NO	ETS-RR-82-24; ETS-TR-82-32
PUB DATE	Nov 82
GRANT	SES80-08481
NOTE	82p.; Some tables may be marginally legible due to small print.
PUB TYPE	Reports - Research/Technical (143) -- Guides - Non-Classroom Use (055)
EDRS PRICE	MF01/PC04 Plus Postage.
DESCRIPTORS	Data Analysis; *Display Aids; Graphs; *Multivariate Analysis; Research Methodology; *Tables (Data); *Visual Aids

ABSTRACT

The past decade has seen a substantial growth in methods and schemes for the display of multivariate data. This paper encompasses a sketch of the history of multivariate displays, from the pre-Columbian Quipu to Chernoff's Face; examines a number of techniques; describes their construction; illustrates their use; and comments on their efficacy. The use of several of these methods is illustrated on social indicators data gathered in the study that Angoff and Mencken did a half century ago in their search for the worst American state. The communicative role of the data table is examined with inside-out plots. Function plots, polygons, tree and face displays are discussed. The issues in "seeing" versus "reading" a display consider Ehrenberg's guidelines. (Author/CM)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

RR-82-24

ED226027

Multivariate Display from Quipus to Faces

Howard Wainer

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

✓ This document has been reproduced as received from the person or organization originating it.
Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
Educational Testing Service
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

TM 830 055

PROGRAM STATISTICS RESEARCH

TECHNICAL REPORT NO. 82-32

EDUCATIONAL TESTING SERVICE
PRINCETON, NEW JERSEY 08541

Multivariate Display from Quipus to Faces

Howard Wainer

Program Statistics Research
Technical Report No. 82-32

Research Report RR-82-24

November 1982

This research was partially supported by the Program Statistics Research Project of the Educational Testing Service, and the National Science Foundation (Grant #SES80-08481) to the Bureau of Social Science Research, Howard Wainer, Principal Investigator.

Copyright © 1982 by Educational Testing Service. All rights reserved.

The Program Statistics Research Technical Report Series is designed to make the working papers of the Research Statistics Group at Educational Testing Service generally available. The series consists of reports by the statisticians in the Research Statistics Group as well as their external and visiting statistical consultants.

Reproduction of any portion of a Program Statistics Research Technical Report requires the written consent of the author(s).

Abstract

The past decade has seen a substantial growth in methods and schemes for the display of multivariate data. This paper encompasses a sketch of the history of multivariate displays, examines a number of techniques, describes their construction, illustrates their use, and comments on their efficacy.

Acknowledgments

I would like to thank Albert D. Biderman for important references, stimulating discussions, and pointed criticism, all of which added considerably to whatever value this essay may have. In addition, this paper has profited from the comments on an earlier draft by: Andrew Ehrenberg, Beat Kleiner, Mary Elsner, Norman Frederikson, and William Angoff.

This paper will be appearing in a Festschrift volume in honor of Herman Chernoff to be published by Academic Press in 1983. All references should be made to this source.

Iconic metaphorical methods for the storage and display of multivariate data have been in use for centuries, from the pre-Columbian Quipu (see Figure 1) to Chernoff's Face (1971) and its subsequent modifications (Wakimoto, 1977; Flury & Riedwyl, 1981). The Quipu is a data display made out of rope that used the number, position, color and type of knot to allow the Peruvian Incans to record a wide variety of statistical data (Locke, 1923). Modern displays employ similar iconic devices, usually, however, on a two-dimensional surface

---

Insert Figure 1 about here

---

rather than hanging freely on a wall or slung over one's shoulder. Despite the apparent variety, the study of the history of multivariate displays reveals more similarities than differences in the aspects of the display characteristics utilized. A major contribution to this ancient methodology was made by Herman Chernoff when he emphasized, through his now famous Face, the usefulness of a display icon which was memorable in its totality. This important characteristic of a display was not emphasized before, and its highlighting by Chernoff has increased our understanding of what makes a useful display.

The development and popularization of the most common of formal methods of data display are generally attributed to William Playfair and his Atlas in 1786. These techniques (the bar, circle and line charts) were generally used for the display of bivariate data (i.e., exports by time), sometimes in combination with other bivariate data (i.e., imports by time), sometimes with the differences between them shaded. The restrictions that plotting on a plane imposed delayed the display of multivariate data in all their complexity for some time.

Insert Figure 2 about here

Nevertheless, techniques did develop, and the reader of 19th Century statistical atlases was treated to a broad array of innovative attempts to display complex multivariate data clearly. The impetus for this seemed to come first from geographers who wanted to show three or more variables--one or more statistical variables on two geographic dimensions. Since the two dimensions of the display plane were used up by the geographic depiction, another visual technique (what Bertin [1973] has since called a 'retinal variable'--as distinct from a locational variable) had to be found to represent the statistical information. The first instance of the shading of a geographic region according to a statistical variable (illiteracy) is credited to Dupin (1819) by Beniger and Robyn (1978). This has become known (pejoratively) as a 'patch map' (Tukey, 1979), which, though it does depict the third variable, can mislead because the visual impact is a function of the size of the area shaded, which may not be related to the relative number of cases involved. This technique has remained in common use. A favorite example, from Fletcher (1849), is shown in Figure 3 depicting the geographic distribution of crime in England and Wales.

Insert Figure 3 about here

One solution to the 'patch map problem' was proposed in 1890 in the Album de statistique graphique de 1889, in which pie charts were placed in each department of a map of France to depict various aspects of the traffic observed there. (See Figure 4.) This notion of using a graphical icon on a geographical

background seems to me to be a powerful idea, and one which provided the inspiration for "Facing the Nation" (Figure 20).

Insert Figure 4 about here

Other graphical icons were developed that could be used to display complex data for a variety of purposes. Often these developments emerged from individuals not commonly associated with statistical display. Florence Nightingale (1857) invented the 'Coxcomb' to show cyclical data. This is a variation on the pie chart theme. In a traditional pie chart (as in Figure 4 above)

Insert Figure 5 about here (Nightingale Rose)

---

the radii of all segments are held constant, and the central angle varies as a function of the data. In the 'Coxcomb' the angle is held constant and the radii are allowed to vary. This visual metaphor is useful when the variable depicted by the particular segments is regular (i.e., in the example above, each segment represents a month, and the area of the segment represents mortality). This approach was used more recently by Fienberg (1975) in an effort to display the contents of contingency tables.

Insert Figure 6 about here

This scheme was used to good effect by Biderman and Neal to display a complex contingency table dealing with data obtained through interviews with rape victims. Aside from details, it shows clearly that in the most dangerous

---

Insert Figure 7 about here

---

situation (being taken away from a public place) the resistance mode adopted seems unrelated to the outcome of the attack.

Another iconic scheme which had its origin in the 19th Century was Anderson's 'metroglyphs' (1957, 1960). These are circles of fixed radii with rays of different lengths representing the values of different variables. A variant of this connects variable length rays originating at the center of the circle to produce irregular polygons (Siegel, Farrell, Goldwyn & Friedman, 1971; Siegel, Farrell, Miller, Goldwyn & Friedman, 1973). The notion is that multivariate points that are similar produce polygons that resemble one another. This same device was independently invented (1857) and published in 1858 as part of the International Statistical Congress in Vienna. Shown in Figures 8 and 9 are versions of these glyphs from this Congress and from the Siegel, et al. paper.

Insert Figures 8 and 9 about here

This technique was well applied by the Biometrics Laboratory of the University of Arkansas (Walls and Epley, 1975) to represent multivariate public health information. Several aspects of this version are worthy of

---

Insert Figure 10 about here

---

notice. First the use of the regular polygon to represent the entire state provides us with a useful referent. Second, the inclusion of a map informs us further of the precise geographic region to which these plots refer. Of course, the graphers assume a certain amount of knowledge on the part of the readers, but this seems plausible, given the prospective audience.

In the search for graphical methods to display multivariate data, the geographers' techniques were adopted and adapted to other fields. Buache (1752) developed the first contour map of physical geography. This technique was widely applied, and in 1843 Lalanne produced a three-dimensional plot of temperature x hour x month (in Figure 11) in which the abscissa was months, the ordinate hour, and the temperature was represented by contour lines. Perozzo (1880) went the obvious next step and developed the stereogram which provided a more visually compelling version of three-dimensional data; Figures 12 and 13 show two versions of these.

Insert Figures 11, 12, and 13 about here

Thus techniques for the display of multivariate displays did develop, and the resulting multiple-bar charts, subdivided bilateral frequency diagrams, perspective drawings, stereograms, contour plots, shaded maps, etc., became familiar to the readers of 19th Century statistical atlases. It is interesting to note that among Francis Galton's other accomplishments was a monograph (1863) describing graphical "methods for mapping the weather," many of which are now the standard. Karl Pearson followed Galton's example and "devoted considerable attention to graphics" (Kruskal, 1975, p. 31). This devotion took a backseat to the development of the mathematical side of statistics during subsequent years, but the past decade has seen a renewed effort in the development of multivariate displays. It would seem that this effort was spurred by the increasing popularity of multivariate statistical procedures and the availability of inexpensive plotters and good software.

Broadly available computing hardware and software made complex analyses easy for everyone, yet their very complexity removed them further from the

intuition of the user. This resulted in the increased possibility of erroneous or inappropriate usage going undetected. Two avenues to guard against this were explored. One was the explosion of interest in robust statistical procedures that yielded protection from many sorts of unforeseen events; the second was the variety of graphical methods that helped to bring into consonance the user's intuition of the multivariate data structure with the results provided by the analytic model.

Many of the techniques developed for looking at multivariate data derived from the same basic notion; an icon of many parts is used to represent a multivariate point, the size and/or shape of each part of the icon represented one component of the data. Herman Chernoff, whom we now honor, was instrumental in this development. His notions about the necessity of a representative icon having a memorable perceptual gestalt joined comfortably with the ideas being expressed by Bertin (1967)[1] regarding the mental process of "visual selection." Together they represent an important contribution to the emerging theory of data display.

In this essay, I would like to discuss and illustrate several multivariate display methodologies. We start our examination of multivariate display methods with the table. Although the table (like the Quipu) has traditionally been a medium for data storage more than display, it can also serve a display purpose. Ehrenberg (1977) has presented a variety of suggestions that allow a table to better communicate the data structure contained in it. Now that data storage

---

[1] It is important to note that this classic work, previously only available in French and German, has recently appeared in an English translation, published by the University of Wisconsin Press.

is primarily accomplished through electronic records, the communicative role of a table must be considered seriously indeed. Among these suggestions are:

1) Massive rounding of table entries. The human eye and mind do not easily perceive differences in numbers with more than two digits. Often rounding to two meaningful digits aids perception enormously.

2) Ordering rows and columns of the table by some aspect of the data. We have previously (Wainer, 1978) referred to this as the principle of, "I'm not interested in Alabama first." Too often we see data tables ordered alphabetically. This aids in locating particular entries for large tables, but does not simplify the perceptual structure of the table. For smallish tables (10-20 rows) alphabetizing is not too crucial. In situations when it is, we can utilize the techniques employed in the 19th Century statistical atlases; the data were ordered by some meaningful aspect, and an alphabetical index to aid in locating entries of interest was provided.

3) Border the table with summary values to aid in comparisons. Adding a robust summary value (a median is usually a good choice) to each column or row (when summarizing is appropriate) facilitates comparisons within that column or row. It also helps to answer preliminary questions that one has when faced with the data.

4) Choose scale wisely. Scale variables so that comparisons that are properly made can be made easily. Those comparisons

that ought not be made, should not be included in the same table.

5) Use space wisely. If some data entries are different from the rest by construction (summary values rather than actual data), physically separate them a bit from the rest. Group together those that are similar. Grouping columns closer together facilitates visual comparisons.

Following these simple and commonsense rules, we will find that the tables thus prepared are easier to understand (and lose little in terms of their accuracy). They will also lead us toward a recently developed multivariate display technique. Rather than deal with display methods in the abstract, we will work within the context of a concrete example. Within this context we will describe several display methods and illustrate them on the same data set. The use of a concrete example has several consequences. It will illustrate the methods, pointing out their relative strengths and weaknesses. It will show how these methods, used in combination, can yield information and insights unavailable from a single method. It will suggest commonalities among the methods that provide the impetus for the development of new techniques for special purposes.

First, the example:

A true multivariate example: The Worst American State--Revisited[2]

> "Just what such words as progress and civilization mean is often disputed, but no one doubts that such things exist. Holland is obviously a more progressive country than Portugal, and equally obviously France is more civilized than Albania. It is when concrete criteria are set up that dispute begins, for every man tries to measure the level of a given culture by his own yardstick...."
>
> (Angoff & Mencken, 1931, p. 1)

---

[2]This section is expanded from Wainer and Thissen, 1981.

Fifty years ago, Angoff and Mencken began their comparative study of social indicators in search of "The Worst American State." Having started out with the basic premise that some places are better to live in than others, they go on to quote Todd that,

> "We shall have to agree that life on the whole is better than death, that health is better than sickness, that freedom is better than slavery, that control over fate is better than ignorance, that moderate provision for human need is better than chronic lack, that broad interests and moderate desires are better than narrowness and enforced asceticism--"

In an effort to reexamine Angoff and Mencken's conclusions half a century later, and to illustrate some multivariate display techniques, we gathered recent data for seven of the same variables from the Statistical Abstract of the United States: 1977.

Tables

Table 1 shows data on ten states for the seven variables indicated. The form of the table is not a recommended one, but is done in a manner that is probably most common. The states are ordered alphabetically; the numbers are given in varying accuracy (population is given to the nearest thousand, life

---

Insert Table 1 about here

---

expectancy to the hundredth of a year [4 days!]). It is clear that although the accuracy represented may be justified by the data, it is not the most useful way to show the data for communicative nor exploratory purposes.

Following Ehrenberg's (1977) suggestions, the data were rounded, states were reordered by an aspect of the data (life expectancy), column medians were calculated and shown, and space was used to block the table, which Ehrenberg (personal communication) suggests facilitates 'reading across.' In addition, the variables were reordered, bringing to the left those variables whose direct

relationship with quality of life was clear (i.e., lower homicide rate is better), and moving to the right those variables whose relationship to quality of life was uncertain (i.e., is it better or worse to have warm weather?).

Insert Table 2 about here

Viewing Table 2 tells us much more than was immediately evident from the original data table--that the two southern states are substantially inferior to the other eight on the first five social indicators. We see that California seems to have a relatively high homicide rate and per capita income. Further we get an impression of the overall levels of each of these indicators from the summaries at the bottom of the table, thus adding meaning to any particular entry. We lose the ability to immediately locate any particular state, which does not seem too important in a short list, but which may become more important with larger tables. Ehrenberg (1977) argues that most tables are looked at by individuals who are not naive with respect to their contents, and so have a reasonably good idea about the approximate position of each member. Thus if one looked at the bottom of a list for a state one expected to find there, and only later discovered that it was somewhere else, this would be useful and surprising information. He claims that the losses associated with not having the table in alphabetical order are modest (for most applications) compared to the gains in potential for internal comparisons.

Table 2 provides us with some insights into the data structure. One of these is that the variables are not oriented in a way that allows easy comparison. For example, we see that when life expectancy is long, homicide rate seems low (not a surprising relationship). Similarly, "% High School Graduates" seems negatively related to the illiteracy rate--again a comforting expected

result. By reordering the variables as well as we can manage so that they 'point in the same direction,' we will gain additional insight. We also make use of this transformation step to make the univariate distributions less skewed. Thus we transform the variables as follows:

Life expectancy--untransformed

Homicide rate (per thousand, denoted P)--transform to non-homicide rate by using $f(P) = -5 \log [P/ (1-P)]$

Income--use square root of income

% High School Graduates--untransformed

Illiteracy rate--use literacy rate = 1 /SQRT (illiteracy rate)

Population--use log(Population) x 10

Days/year below freezing--untransformed

The resulting table (Table 3) is not as close to intuition as Table 2 because of the transformations, but is a useful next analysis step. While the

---

Insert Table 3 about here

---

transformed numbers may be further from intuition than the originals, we can see the same structure that was evident in Table 2, and the problem of scale becomes clearer still. In a true multivariate situation, like this one, making comparisons across variables is difficult. Some transformation is required. In the original work Angoff and Mencken used order statistics, which seems reasonable, but we can do better.

Inside-Out Plots

The first step toward the alleviation of the problem of comparing across variables is to center the table by column. We can do this easily and robustly by subtracting out column medians and displaying the residuals. This is shown

in Table 4. We note more clearly that the bottom of the table tends to have many negative residuals, the top many positive ones. This reassures us that the variables have been oriented correctly, and that they do, in fact, form a positive manifold in the variable space, with the possible exceptions of population and temperature.

Insert Table 4 about here

Even after column-centering the table, we see that the residuals are still quite different and require some rescaling. An easy and relatively robust way to do this is to divide each column by the median of the absolute value of each residual in that column. This is the well-known MAD estimator of spread. When we do this (Table 5), we have finally arrived at a table that, although the numbers are removed from their original meaning (i.e., an entry of .3 under 'coldness' for Iowa is a long way from the original '140 days a year below freezing'), they now allow us to compare the relative standing of each state across variables. This was not possible in the original metric. If the aim of the original table was not to facilitate these comparisons, why were these data included in the table in the first place?

Insert Table 5 about here

Now that the states are comparable across variables (the entries can be thought of as proportional to robust z-scores), we can row-center the table by calculating and subtracting out row medians. These medians are shown flanking Table 5 and provide one measure to answer the original question, "What is the worst American state?" Our answer, based upon only these 10

states, matches the answer that Angoff and Mencken found half a century ago. In their words,

> "On all of these counts Mississippi fails. It has few natural resources, and suffers from a bad climate and a backward population. It produces only a fourth as many candidates for "Who's Who in America" relative to its population, as Massachusetts. It has no efficient police, as its lynching record shows, and its government is in the hands of office-seekers of low character. It is also deficient in decent hospitals, colleges, newspapers and libraries" (p. 371).

Although our study did not include as many variables as did Angoff and Mencken's, the position of Mississippi is identical--last. We can see the relative position of the various states rather well by their ordering and spacing in Table 5. If we use space more conspicuously to relate the relative position of the ten states overall, we arrive at the diagram shown in Figure 14. This is a back-to-back stem-and-leaf display (Tukey, 1977) in which the row medians are written down linearly in the center (the stem), and the states which have these medians are written down adjacent to the appropriate median (the leaves)[3]. This can be usefully thought of as turning a table inside out, for we ordinarily have labels on the OUTSIDE of a table and numbers on the INSIDE. For a stem-and-leaf display we move the numbers outside to make them labels, and the labels inside next to the appropriate number. This concept of inside-out plotting will be even more useful shortly.

---

[3] The values plotted here were arrived at with somewhat more precision than what was shown in Table 5, thus the slight discrepancies. Obviously, our advocacy of extreme rounding for presentational purposes does not extend itself ordinarily into intermediate calculations.

Insert Figure 14 about here

Figure 14 shows the overall position of each of the ten states. What remains when the data matrix in Table 5 is row-centered by subtracting out these overall positions is what is of interest next. Table 6 shows the final matrix. It has been column-centered by removing column medians, column-scaled by dividing each column by the median absolute deviation from the column median, and last, row-centered by subtracting out the row medians. All of this was done after the variables were oriented in the same direction and transformed to be as symmetric as possible (within generous limits). The resulting residual matrix is now ready for display.

Insert Table 6 about here

Glancing over Table 6 our eyes note several unusually large entries. It is hard to see much order in the residuals. Can we display this in a way that makes the features of interest more obvious? Ramsay (1980) has suggested that plotting Table 6 inside-out is a way that answers this question straightaway. To do this, we prepare stem-and-leaf diagrams of each column of this table separately, but since they are on comparable scales we can plot them side-by-side on a common stem (see Figure 15). Thus turning this table inside out--placing

---
Insert Figure 15 about here
---

the labels inside and the numbers outside--we immediately see details of structure that add to the states' overall ranking. We see that Pennsylvania seems to have a shorter life expectancy than its overall rank would predict--perhaps looking at infant mortality rates and black lung statistics would be

helpful to understand this. We note that California's literacy rate seems unusually low when one considers its high overall ranking, suggesting an interpretation based upon its large Spanish speaking population.

The plotting of multivariate residuals inside-out (as developed by Ramsay) follows logically from the basic premises of tabular display espoused by Ehrenberg. They allow the tabular format to communicate information and provide that most valuable of graphic services--"force us to notice what we never expected" (Tukey, 1977). We discover that many of the findings seen clearly in Figure 15 were visible in Table 2, now that we know to look for them.

This display has the added advantage of getting better as the data set (number of observations) gets larger. We are typically only interested in observations with large residuals; therefore when the data bunch up in the middle (as would be the case with a larger data set), we can safely ignore them. Further, this display method works well when data are not scored in the same way, though it does take some preprocessing to get the data ready for plotting. That includes a variety of two-way scatter plots to determine the best orientation of each variable. In this instance we ordered 'number of days below freezing' in the same direction as literacy not because of any love for cold weather, but rather because it was in this direction that it related to all the variables that are easily oriented. (Angoff and Mencken ordered 'number of Baptists' as 'more = worse' partially because of its high correlation with 'number of lynchings' in that orientation.)

The operations described previously involved with the transformation of variables prior to, or in the midst of, preparing a multivariate display are common to all display methodologies. One must almost always make a sequence

of displays--each one telling how the next should be scaled. Univariate stem-and-leaf diagrams were used to help determine the transformations needed for symmetry. Bivariate plots aid in determining the correct orientation of variables and point toward unusual points that might confuse solely analytic transformation methods. However, while the analyses are done sequentially, the wise analyst often returns to previous displays with newfound knowledge, to redisplay the data after subsequent analyses pointed toward an important or useful transformation. In addition to providing additional information about the data set, this technique also reduces the effect that the particular sequence used has had on the results.

Function Plots

Andrews (1972) developed a scheme for multivariate display that is quite different from most methods used previously. His method involves calculating a periodic function of k-Fourier components for each k-dimensional data point. The value of each Fourier component of that function is determined by the value of its associated variable. The strength of this technique is that it allows the inclusion of many variables but limits the number of data points tnat can be effectively shown; with more than 10 or 20 points the plot can get too busy to be helpful. Another disadvantage is that such plots are quite far removed from the data, so that even though one can see which points tend to group together, one must look to other display methods to better understand why.

Shown in Figure 16 is a function plot of the Worst American State data. We note the similarity of Alabama and Mississippi and the close matching of Pennsylvania and Ohio. If one chooses the plots with care, such characteristics

---

Insert Figure 16 about here

---

as 'overall height' and 'general jiggliness' can have particular importance.

As with almost any innovative methodology, experience with its use is required before the subtleties of its character reveal themselves. The potential value of function plots does not always manifest itself upon first acquaintance. Tales are told in the hallways of Bell Laboratories about a contest to discover the secret of a data set constructed especially for this purpose. There several of the talented statisticians examined the problem under the illumination of their favorite display/analysis methods. The only one to come up with the correct solution was David Andrews, who used function plots. The apportioning of credit for this victory between the tool and the craftsman is difficult, yet it does supply support for the inclusion of function plots within the toolbox of the serious data analyst.

Polygons

A more traditional icon for multivariate display is a polygon formed by k-secting a circle with k-radii each associated with a particular variable. The length of each radius is proportional to the value of the variable for the data point to be displayed, and these various points are then connected to yield an irregularly shaped polygon. This technique is more than a century old and provides an often useful display. (See Figure 8.) Usually each variable is scaled separately, so that the maximum value of each variable touches the circle. Sometimes this is varied to make a regular polygon with all radii touching the circle a standard (i.e., this could represent the condition for the entire U.S.). Then any point that stuck out past the circle's perimeter would represent 'greater than average.' We chose the former convention in our display of the Worst American State data shown in Figure 17. Note the convention of orienting all variables so that 'bigger = better' is quite useful. Thus states that are in the running for the title of 'Worst' have smallish

polygons, whereas those at the other end look much larger. Note further that it is relatively easy to discover (using the legend) what are each state's strong and weak points. The shortcomings of this method are:

1) Producing a legible display with many more than seven variables becomes difficult.

2) The shape of the polygons is arbitrary, because of the arbitrariness of the order of the variables around the circle.

3) Odd shaped polygons are not memorable, so one is hard pressed to remember any details about a state other than its gross structure (if it was small or large, regular or irregular--this latter bit of structure can be varied by reordering the variables).

4) The legend is crucial to the display. It seems difficult to keep in mind which corner refers to which variable. Take away the legend and the interpretability of the display shrinks.

5) Similarities in shape seem to be perceived irrespective of differences in orientation. Thus Vermont and Oregon seem similar (except for rotation). Of course, this similarity is to some extent false, due just to the particular ordering of the variables.

Despite these shortcomings, the polygon can be a useful display.

Insert Figure 17 about here

Returning to our investigation into the Worst American State and Figure 17, we see: Alabama and Mississippi standing out, the similarities between Ohio and Pennsylvania (among the causes of this similarity is the relatively low life expectancy), and Iowa emerging as a contender for the other end of the spectrum.

Trees

In the display methods so far discussed the viewer is able, more or less well, to visually cluster the multivariate observations (states) through of the similarity of their representative icons. Clustering of the variables (state characteristics) is not easily done. This inability to represent simultaneously the clustering of variables and of observations I view as a weakness of the display. Kleiner and Hartigan (1981) solved this problem in a simple and ingenious way. They noted that a tree is a common icon to represent a hierarchical cluster structure, and that the structure of the tree is reasonably well determined (Johnson, 1967; Gruvaeus and Wainer, 1972). They reasoned that one could use a variation of the tree structure obtained from the inter-observation distances as a basic icon. This single tree would provide the shape of the icon, but the size of each branch of the tree would be determined by the value of the variables represented by that branch. Thus each observation would be represented by a tree with the same general structure, but with different size branches.

In Figure 18 we see a tree derived from the Worst American State data (averaged across all ten states). We see immediately that there are two major groupings of variables: one includes life expectancy, percent of high school graduates and income; and the other includes homicide rate, literacy and temperature. The populations of the states appear to be a relatively isolated

variable. In Figure 19 this tree is distorted to reflect each state's multi-

---

Insert Figures 18 and 19 about here

---

variate structure. We note California's asymmetry with warm weather and high homicide rate shrinking the left cluster; the symmetric structure of Vermont, New Hampshire and South Dakota is similar to that of Iowa, but their small populations separate them. Mississippi and Alabama are, again, shown to be distinctive.

We can conclude that the tabular scheme seems better at supplying details, for clustering states function plots win out, but clustering of variables is done well only by trees. Polygons seem to provide adequate performance in displaying qualitative detail and clustering observations, and so is a reasonable compromise candidate. It is also easier to use within the context of mass communication than the other three methods. It provides a more eye-catching image than a table and is easier to explain than a tree or function plot. Inside-out plots are a curious mixture of tabular and graphic display. They provide a clear view of the best unidimensional ordering, as well as highlighting deviations from this order.

Faces

I previously referred to the polygon display as "a more traditional icon." Let me take this opportunity to explain that remark a bit further, and thus better set the stage for the discussion of the Chernoff Face. A survey of multivariate icons reveals that many of them share a common structure. Typically, each icon is made up of many parts, each of which varies in size and shape as a function of the data. Thus the final appearance of the whole icon reflects contributions of the various parts. Polygons obviously fit

this mold, as do Anderson's (1960) metroglyphs, Bertin's (1977) graphical matrices, Kleiner and Hartigan's (1981) trees, as well as more ancient types like the quipu and the various visual metaphors employed by cartographers. To some extent so too does the Chernoff Face, with an important variant. In the Face construction the size, shape and orientation of each feature is related to a characteristic of the data, yet the icon is conceived of as a whole and is intended to be perceived from the outset as a gestalt--not as the sum of its parts.

Before we discuss this icon in greater detail, let us first look at its use within the context of our example--the Worst American State. The first step in displaying data with faces (after the sorts of orientations and scalings described previously) is the assignment of variable to feature. This is not always easy nor is it unimportant. Jacob (1978; 1981) has suggested several empirical aids to this assignment problem; they supplement but do not supplant displayer wisdom. The assignment plan we arrived at (after several tries) was:

1) Population => the number of faces/state--The number of faces is proportional to the log of the population. We used one large face in each state for easier identification and as many identical small faces as required. The size of these latter faces was kept small so as to allow us to fit them within the confines of the state boundaries.
2) Literacy rate => size of the eyes (bigger = higher)
3) % HS graduates => slant of the eyes (the more slanted the better)

4) Life expectancy => the length of the mouth (the longer the better)

5) Homicide rate => the width of the nose (the wider the nose, the lower the homicide rate)

6) Income => the curvature of the mouth (the bigger the smile, the higher the income)

7) Temperature => the shape of the face (the more like a peanut, the warmer; the more like a football, the colder)

8 & 9) Longitude and latitude => the X and Y position of the face on the coordinate axes of the paper represents the position of the state.

Thus we tried to use sensible visual metaphors for representing each variable; 'Bigger = Better' was the general rule when normative direction was clear. In the case of a variable (such as weather) where desirability could not be determined easily, we used an aspect of the face that is not ordered. To show the adaptability of the FACE scheme to larger data sets, we prepared a plot involving all fifty states. Similar plots could have been prepared for most other meta-iconic schemes (STARs, TREEs, etc.). The resulting plot is shown in Figure 20.

Insert Figure 20 about here

A viewing of the map reveals many things. First, we can see the relative density of population in the East at a glance. Next, note the temperature gradient as one goes south from North Dakota to Texas. The Pacific current is also evident. Looking more closely we see that the deep South looks very

homogeneous and low on all variables of quality of life used. Thus we can see that Angoff and Mencken's finding that:

> "The Cotton Belt...is the least advanced part of the United States, and of all the cotton states Mississippi is the most unfortunate, for it lacks the coal and iron of Alabama, the oil of Texas and Arkansas, the trading opportunities of Louisiana, and the manufacturing capabilities of Georgia and the Carolinas. In the midst of its hordes of barbaric peasants there is some native stock of excellent blood. But the young men of this stock, finding few opportunities at home, have to go elsewhere. Altogether, it seems to be without a serious rival to the lamentable preeminence of the Worst American State" p. 371.

is not disconfirmed by our reanalysis of similar data taken half a century later (at least using these variables).

Note further that three New England states, Vermont, New Hampshire and Maine, look the same: they are generally high on education variables, as well as having low homicide rates, but appear to have low incomes. Massachusetts seems to have a somewhat lower literacy rate but higher per capita income. Connecticut and New Jersey seem remarkably similar with higher incomes still. We also see a clustering of Mid-East states on all variables (Pennsylvania, Ohio, Michigan and Indiana) with the border states of Virginia, West Virginia and Kentucky falling somewhere between their rural neighbors and their more industrial ones.

The Mid-West looks homogeneous with the rather surprising difference between North and South Dakota in income. A check back to the original data indicates that this is not a data entry error. There is a substantial difference between these two neighbors. We also note that except for the cold weather, North Dakota seems a pleasant place to live. Utah has an interesting structure, being very high on all variables except income, which may reflect the influence of the Mormon church on the state. There are many other interesting

similarities that might be noted, and the reader is invited to study the map more closely. As a last comment, we draw the reader's attention to the similarity between California and Hawaii. We see benign weather, long life, high income, low homicide rate, high proportion of high school graduates, but a low literacy rate. This reflects either their high proportion of non-English speaking inhabitants, or a propensity for lotus eating so common in paradise.

The map in Figure 20--Facing the Nation--has several unique characteristics. The human face quality of the display icon makes the display memorable. Without having the map in front of you, it is easy to characterize from memory the sad squinty-eyed face of the South. Similarly, the broad faced stolid Mid-Westerner emerges clearly from memory, as do the smiling faces of Hawaii and California. Once one remembers the correspondences between facial features and statistical variables, the attaching of meaning to the visual memory becomes easy. Quantitative details are lost, but general structure remains. Visual clustering in k-space is done almost without noticing. Care must be taken in the assignation of features, for some seem to have a more profound effect upon perception than others. Yet, with the addition of a little wisdom a useful display appears.

Other Uses and Extensions

An instructive example of the use of a face-like display in addition to (or instead of) a more traditional multivariate display was provided by Thissen, Baker and Wainer (1981). The problem they examined was depicting the influence associated with various points in a scatter plot. Devlin, Gnanadesikan, and Kettenring (1975) advocate the use of a scatter plot enhanced with influence contours to aid in the detection of critical outliers. An example of this is

shown in Figure 21. This is a useful plot and allows the easy identification of so-called 'leverage points' (points that exert an unusual amount of influence on the value of the statistic being calculated--in this case the sample correlation coefficient). Thissen et al. point out that for samples of modest size the calculation of the influence contours is much more difficult than simply calculating the influence of each point separately. They go on to illustrate this graphically by plotting a circle for each point, whose area is proportional to the influence of that point. This is shown in Figure 22. This plot does not show the direction of the influence, only its absolute amount. To incorporate the sign of the influence, a curved line (concave up or down) was included in the circle. The result of this is reproduced as Figure 23--the eyes are whimsical and serve to make the image more memorable.

Insert Figures 21, 22, and 23 about here

Chernoff's original formulation has been adapted, modified and improved upon. Flury and Riedwyl (1981) developed a modified face that is a more realistic version with much greater detail (see Figure 24). There is less implicit covariation among the features than in the Chernoff version, and the two sides of the face can have features assigned to them independently, thus allowing the possibility of an asymmetric face.

In an experiment that Flury and Riedwyl ran, they found that their version of the face (used either symmetrically or asymmetrically) provided for more accurate judgments as to which pairs of twins were monozygotic or dizygotic.

Wakimoto (1977) in a further extension of the Chernoff idea used an entire body plot to represent various physiological measurements (see Figure 25). This may be a useful form for certain kinds of applications, but aside from providing

room for a few more parameters, does not seem to improve on the original formulation. It is just another realistic icon.

Insert Figures 24 and 25 about here

This leads to an interesting question. Wakimoto's body plots use the characteristics of the body to represent the physiological characteristics. Thus a fat figure implies an overweight person, small eyes represent poor eyesight, a tall figure represents a tall person, etc. This means that we have an excellent visual metaphor, with the legend able to be learned with little effort. Should this concept be generalized? If we want to represent bank notes (as in Flury and Riedwyl), why use faces? Why not cartoon bank notes with the features measured enhanced in some way? If we want to represent various characteristics of housing, why not use cartoon houses? Previously, I suggested (Wainer, 1979) that an icon like the Wabbit (see Figure 26) be used on cereal boxes to convey nutrient information to children. In this instance the characteristics of the Wabbit's face would be related to the vitamins whose deficiency affects them.

Insert Figure 26 about here

Obviously, such methods have much to recommend them, yet they are not practical as a general approach. The programming necessary to draw a specialized icon for a single application is too costly in time and resources to justify itself unless this use is to be routinized. Thus the power of the face as a generalizable multivariate display--it provides a memorable display for data from a variety of origins. It may not be optimal in any given circumstance, but may provide (as Mosteller & Tukey [1968] describe the

Jackknife--another general technique) 'the broad usefulness of a technique as a substitute for specialized tools that may not be available' (p. 134).

Conclusions

In this essay we have tried to sketch some of the rich history associated with the development of methods for the display of multivariate data. In the course of this sketch we focussed on a few methods that are either exemplary of a class of others or are sufficiently useful and unique to warrant inclusion. The use of several of these methods was illustrated on social indicators data gathered to compare the standings of states in modern America with the study that Angoff and Mencken did half a century ago in their search for the Worst American State. Incidently, we discovered their conclusion to be true that the Cotton Belt states in general, and Mississippi in particular, still appear to deserve that unappealing appelation.

To summarize what was illustrated in this essay we noted:

1) The table did a pretty good job once it was redone, although the full (50 x 7) data set for all states would have been less transparent. Enhanced tables, from which main effects have been removed do better. Still better for large tables are inside-out plots which convey the detail of a table along with a more visually compelling image of the distributions and clusterings of the data.

2) Function plots are sufficiently far from convention that considerable user experience seems to be required before being able to fully utilize the information carried in them. They do seem to easily provide clustering information, although they can be limited in the number of data points

that they can convey and still allow the extrication of identification information of individual points. Their strength is the number of variables that can be portrayed.

3) Polygon plots, as well as other sorts of iconic displays based upon non-metaphorical icons, are useful for qualitative conveyance of information and are made more useful still if the icons are displayed in a position that is meaningful. This was illustrated in Figure 4, and more evocatively still in Figure 20, when the geographic location provided information about the identification of each icon that was far more visually compelling than merely tagging each with the associated state's name. This led us to an important conclusion about multivariate plotting methods--do not waste the two dimensions of the plane. Early plotting schemes showed just two variables by using the two dimensions of the plane to represent them. Multivariate display methods allow many variables to be shown, yet we would be wise not to waste the physical dimensions, for they provide one of the most compelling of visual metaphors.

4) Trees are the only display method that do a good job at depicting the covariance structure among the variables. Their perceptural complexity diminishes their usefulness as a memorable icon, but they are a new idea--still under development.

5) Faces (and subsequent versions including nicer faces, bodies,

and wabbits) allow the display of complex data in a memorable way. There are some problems with their use, particularly those associated with the assignation of features. This problem is eased through the use of clustering schemes like those used by Kleiner and Hartigan (1981) in the formation of their Trees and Castles. Further details of this point are contained in Jacobs' (1981) discussion of their paper.

Seeing vs. Reading a Display

A recurring issue in discussions of display methodologies relates to the kinds of cognitive processes that ought to be brought into play when confronted with a display. Bertin (1977) contends that a display is most efficient when we can "see" the data structure. This entails the use of a display in which the data are represented by a compelling visual metaphor. Shading (darker = more) and stereograms (higher = more) are two examples of such metaphors. Displays that are less efficient need to be "read." Colors used for shading often need to be read (i.e., red = most, blue = medium, yellow = least). Obviously, when a display needs to be read, we both perceive less efficiently and remember more poorly. The great power of Chernoff's invention is that it allows the seeing of quite complex data structures.

Tables can be made more 'seeable' through Ehrenberg's guidelines, but this rests to some extent on the iconic characteristics of the numbers themselves. It is apparent that two-digit numbers look bigger than one-digit ones, and thus we can 'see' that 123 is bigger than 62 and is greater than 8. Confusion manifests itself when the numbers being compared are of the same order. Here we are helped because the numbers themselves are somewhat ordered--1 looks smaller than 2, and 3 looks smaller than 8. This relationship is not perfect--7 looks

smaller than 6--yet it does help. Bachi's (1968) Graphical Rational Patterns are an attempt to make the numbers more graphical. This same conception is behind tallying schemes (/, //, ///, ...) as well as quipu accounting. Thus though numerical representation can be made somewhat graphical and hence 'seeable,' it is important that we are aware of their limitations.

If a display must be read, is that bad? This question seems to be often misunderstood. I believe that a graphic display can serve many purposes. Sometimes we want a display to show general structure quickly and efficiently, and provide a memorable image that can subsequently be retrieved and utilized. This certainly is one of the more important uses of graphical display, and one on which we have concentrated most of our attention. There are others. We note that communication with prose tends to be linear--the storyteller starts at the beginning and proceeds to the end in a manner that appears to him to be logical. Yet stories, or even 'how-to' directions can be told in many ways. If one was to prepare directions as to how to get up in the morning and get to work successfully, there are many possible orders that could be specified--but only one would be. Such is the nature of linear presentation. John Tukey acknowledges this in his classic text (1977) when he points out alternative orders in which the book can be profitably read--some chapters are in series, others in parallel.

A powerful use of a graphic display is to present information in a non-linear way. Thus exploded diagrams of automobile transmissions show clearly which pieces go where and indicate clearly which orders of assembly are possible and which are not. Similarly, the complex charts of population often found in statistical atlases provide many stories of immigration trends. These can be read and studied from left to right, top to bottom, or in many

other nonlinear ways and provide fresh insights. I do not mean to imply that such complex charts are superior to telling any particular tale--rather they are a reasonably compact way of providing the raw material for many stories. Such displays have their place and should not be denigrated when this is the purpose for which they are employed. Our primary concern in this essay is with displays that aid in seeing structure, but we would have been remiss in not mentioning this other role.

An Appreciation

A glance over the history of display indicates the many contributors who have conceived important and useful methods to allow us to better look at our data. I am indeed grateful to Herman Chernoff who has allowed our data to look back at us.

References

Anderson, E. (1957) A semigraphical method for the analysis of complex problems. Proceedings of the National Academy of Sciences, 13: 923-927; reprinted in Technometrics, 2: 3. 387-392 (1960).

Andrews, D. F. (1972) Plots of high-dimensional data. Biometrics, 28: 125-136.

Angoff, C., and Mencken, H. L. (1931) The worst American state. American Mercury, 31: 1-16, 175-188, 355-371.

Bachi, R. (1968) Graphical Rational Patterns: A New Approach to Graphical Presentation of Statistics, Jerusalem, Israel, University Press.

Beniger, J. R., and Robyn, D. L. (1978) Quantitative graphics in statistics: A brief history. The American Statistician, 32: 1. 1-11.

Bertin, J. (1977) La graphique et le traitement graphique de l'information. Paris: Flammarion.

Bertin, J. (1973) Semiologie Graphique. The Hague: Mouton-Gautier. 2nd ed.

Bertin, T. (1967) Semiologie Graphique. Paris: Gouthier-Villors.

Biderman, A. D. and Neal, D. (1979) How to resist rape attacks. The Newsletter of the Bureau of Social Science Research, 8: 9.

Buache, P. (1752) Essai de geographie physique. Memoires de l'Academie Royale des Sciences, 399-416.

Chernoff, H. (1971) The use of faces to represent points in n-Dimensional Space Graphically, Technical Report No. 71, Stanford University, Department of Statistics.

Devlin, S., Gnanadesikan, R., and Kettenring, J. (1975) Robust estimation and outlier detection with correlation coefficients. Biometrika, 62: 531-545.

Dupin, C. (1819) Carte de la France Eclairea et de la France Obscure, Paris.

Ehrenberg, A. S. C. (1977) Rudiments of numeracy. Journal of the Royal Statistical Society, A 140: 277-297.

Fienberg, S. E. (1979) Graphical methods in statistics. The American Statistician, 33: 4. 165-178.

Fienberg, S. E. (1975) Perspectus Canada as a social report. Social Indicators Research, 2: 153-174.

Fletcher, J. (1849) Moral and educational statistics of England and Wales. Journal of the Statistical Society of London, 12: 151-176, 189-335.

Flury, B., and Riedwyl, H. (1981) Graphical representations of multivariate data by means of asymmetrical faces. Journal of the American Statistical Association, 76: 376. 757-765.

Galton, F. (1863) The weather of a large part of Europe, during the month of December 1861. Meteorographica or Methods of Mapping the Weather, London and Cambridge.

Gruvaeus, G. T., and Wainer, H. (1972) Two additions to hierarchical cluster analysis. British Journal of Mathematical Statistical Psychology, 25: 200-206.

Jacob, R. J. K. (1978) Facial representation of multivariate data. In Graphical Representation of Multivariate Data, (ed.) P. C. C. Wang, New York: Academic Press, 143-168.

Jacob, R. J. K. (1981) Comment on Trees and Castles. Journal of the American Statistical Association, 76: 374, 270-272.

Johnson, S. C. (1967) Hierarchical Clustering Schemes. Psychometrika, 32: 241-254.

Kleiner, B., and Hartigan, J. A. (1981) Representing points in many dimensions by trees and castles. Journal of the American Statistical Association, 76: 374. 260-269.

Kruskal, W. H. (1975) Visions of maps and graphs. In Auto Carto II: Proceedings International Symposium Computer-Assisted Cartography, Washington, DC: Census Bureau.

Lalanne, L. (1845) Memoire sur les tables graphiques et sur la geometric anomorphique appliquees a diverses questions qui se Rattochent a l'art de l'engenieur, Paris: read 1843.

Locke, L. L. (1923) The ancient quipu or Peruvian knot. In The American Museum of Natural History: Science Education, Library of Congress. Smithsonian, 12-73.

Mosteller, F., and Tukey, J. W. (1968) Data analysis: Including statistics. In The Handbook of Social Psychology, (eds.) G. Lindsey, and E. Aronson, 80-203, Reading, Mass: Addison-Wesley.

Nightingale, F. (1857) Mortality of the British Army, London.

Perozzo, L. (1880) Della rappresentazione graphica di una collettivita di individui nella successione del tempo. Annali di Statistica, 12: 1-16.

Playfair, W. (1786) The Commercial and Political Atlas, London.

Ramsay, J. O. (1980) Inside-out displays and more. Presented at a Symposium on multivariate data displays. Psychometric Society Meeting, Iowa City.

Siegel, J. H., Farrell, E. J., Miller, M., Goldwyn, R. M., and Friedman, H. P., (1973) Cardiorespiratory interactions as determinants of surgical and the need for respiratory support in human shock states. J. Trauma 13: 602.

Siegel, J. H., Goldwyn, R. M., and Friedman, H. P. (1971) Pattern and process of the evolution of human septic shock, Surgery, 70: 232-245.

Thissen, D., Baker, L., and Wainer, H. (1981) Influence-enhanced scatterplots. Psychological Bulletin, 90: 1. 179-184.

Tukey, J. W. (1979) Methodology and the statistician's response for both accuracy and relevance, Journal of the American Statistical Association, 74, 368, 786-793.

Tukey, J. W. (1977) Exploratory Data Analysis, Reading, Mass: Addison-Wesley.

Wainer, H. (1979) The Wabbit: An alternative icon for multivariate data display. BSSR Technical Report, 547-792.

Wainer, H. (1978) Graphical display. Presented at European Mathematical Psychological Association, Uppsala, Sweden.

Wainer, H., and Thissen, D. (1981) Graphical data analysis. Annual Review of Psychology, 32: 191-241.

Wakimoto, K. (1977) A trial of modification of the face graph proposed by Chernoff: Body Graph. Quantitative Behavior Science (Kodo Keiryogaku), 4: 67-73.

Walls, R. C., and Epley, E. A. (1975) Arkansas Health Graphics, 15-135 Department of Biometry: University of Arkansas.

Figure Captions

Figure 1 An example of the Incan Quipu. The pendant strands are grouped in fours, each group being tied with a top strand. The top strand sums the numbers on the pendant strands (Locke, 1923).

Figure 2 An early line graph (originally in color) done by Playfair (1786).

Figure 3 One of many shaded maps done by Fletcher (1849) to depict the geographic distribution of various social indicators. In this case crime in England and Wales.

Figure 4 An early attempt to display several statistical variables on a geographic background (Album de statistique graphique de 1889).

Figure 5 One of Florence Nightingale's (1857) 'coxcomb' displays, showing monthly casualties during the Crimean War. (Originally in color.)

Figure 6 A modern variant of Nightingale's design by Fienberg (1975) to show a 2-by-2 contingency table.

Figure 7 An application of Fienberg's scheme used by Biderman and Neal (1979) to show the multivariate results of a study of the effects of various resistance strategies in the crime of rape.

Figure 8 Polygon plot from Siegel et al. (1979). At corners are shown prototype patterns for the A, B, C, and D states. In the large circle in the center is the physiological pattern manifested by a 44-year-old woman. The circle in the center represents the reference control state. Each dotted circle represents one standard deviation from the reference state.

Figure 9 Early example of polygon plots from the 1857 International Statistical Congress in Vienna.

Figure 10 Polygon plots from Walls and Epley (1975).

Figure 11 Three-way table (temperature x hour x month) reconstructed as a contour map, the first such application of this form. Invented by Lalanne in 1843. Published in 1845.

Figure 12 Relief drawing (stereogram) representing the population of Sweden, 1750-1875, by age groupings, constructed by Luigi Perozzo, Italian Department of Statistics, in 1879. Published in 1880.

Figure 13 Another Perozzo stereogram published in 1880. This one shows the probability of getting married as a function of the ages of the husbands- and brides-to-be.

Figure 14 Back-to-back stem-and-leaf display of the Worst American State effects showing both state names and their two-letter abbreviations.

Figure 15 Inside-out display of residuals in Table 6. Plot allows viewer to center attention on unusual state-variable interactions by showing unusual residuals.

Figure 16 Function plot of Worst American State data.

Figure 17 Polygon plot of Worst American State data.

Figure 18 Cluster diagram of the variables considered in the search for the Worst American State. This tree forms the basic template for the trees in Figure 19 (plot prepared by Beat Kleiner).

Figure 19 Tree display of Worst American State data (plot prepared by Beat Kleiner).

Figure 20 The Worst American State data shown for all 50 states as Chernoff Faces.

Figure 21 Scatterplot of 32 observations (r = .05) with influence contours for influence of .1 and .2 superimposed.

Figure 22 Scatterplot of data of Figure 21, showing circles with areas that are proportional to their influence plotted instead of points.

Figure 23 Scatterplot of the data of Figure 21 displaying schematic faces (with area equal to influence) smiling for positive influence and frowning for negative influence.

Figure 24 Two Flury and Riedwyl asymmetric faces plotted symmetrically. One face shows each feature at its minimum value, the other shows all features at their maximum.

Figure 25 Wakimoto's Deshi, which uses a whole body to convey multivariate information.

Figure 26 Wainer's Wabbits shown here in several variations. These were originally recommended as a display icon to communicate multivariate nutritional information to children.

Figure 1

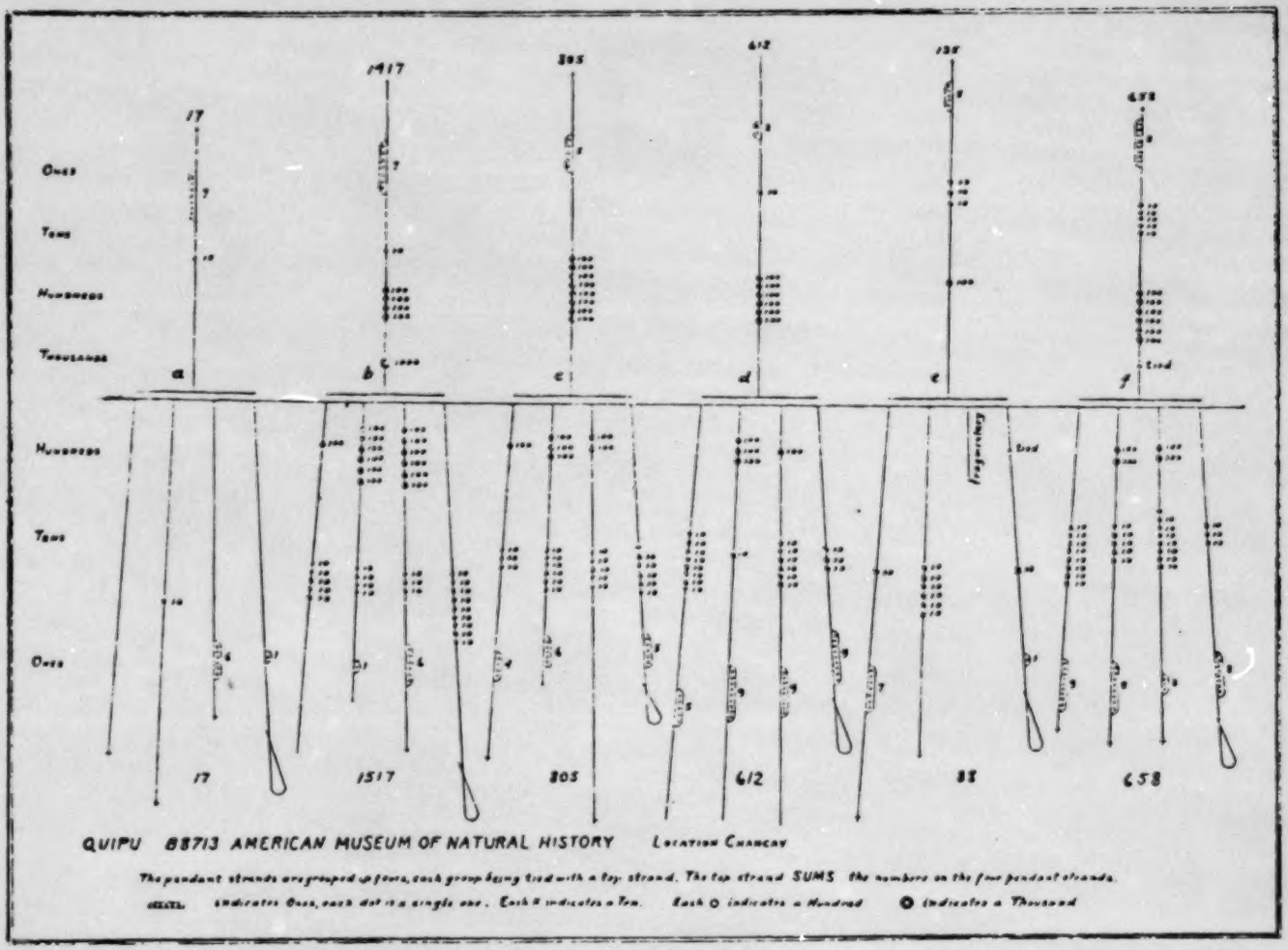

17
1417
805
612
88
658
Ones
Tens
Hundreds
Thousands
a
b
c
d
e
f
Fringe
17
1517
805
612
88
658
QUIPU 88713 AMERICAN MUSEUM OF NATURAL HISTORY Location Chancay
The pendant strands are grouped in fours, each group being tied with a top strand. The top strand SUMS the numbers on the four pendant strands.
indicates Ones, each dot is a single one. Each indicates a Ten. Each ○ indicates a Hundred ● indicates a Thousand

Figure 2

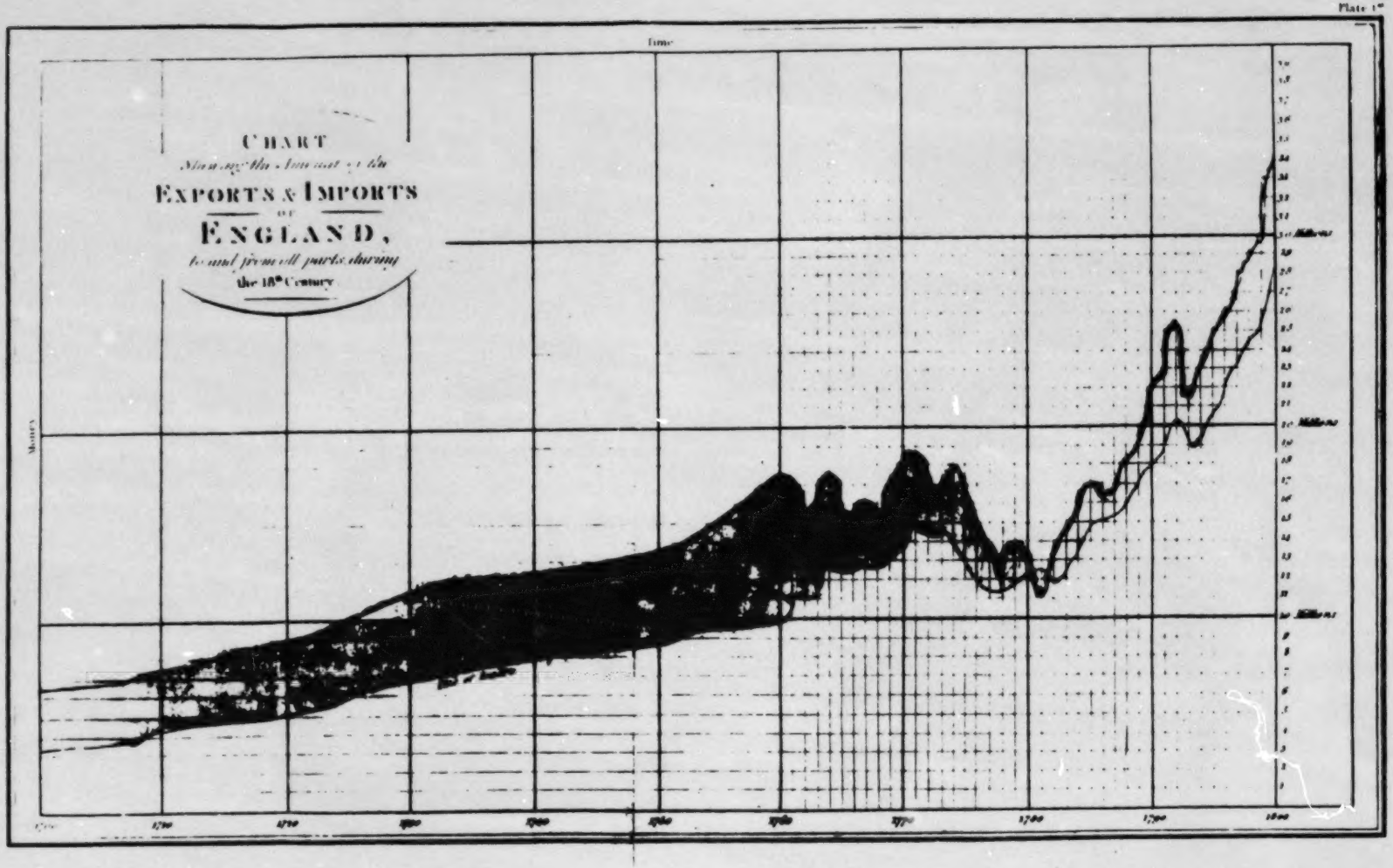


Plate 1st
Time
Money
CHART
EXPORTS & IMPORTS
OF
ENGLAND,
the 18th Century

Figure 3

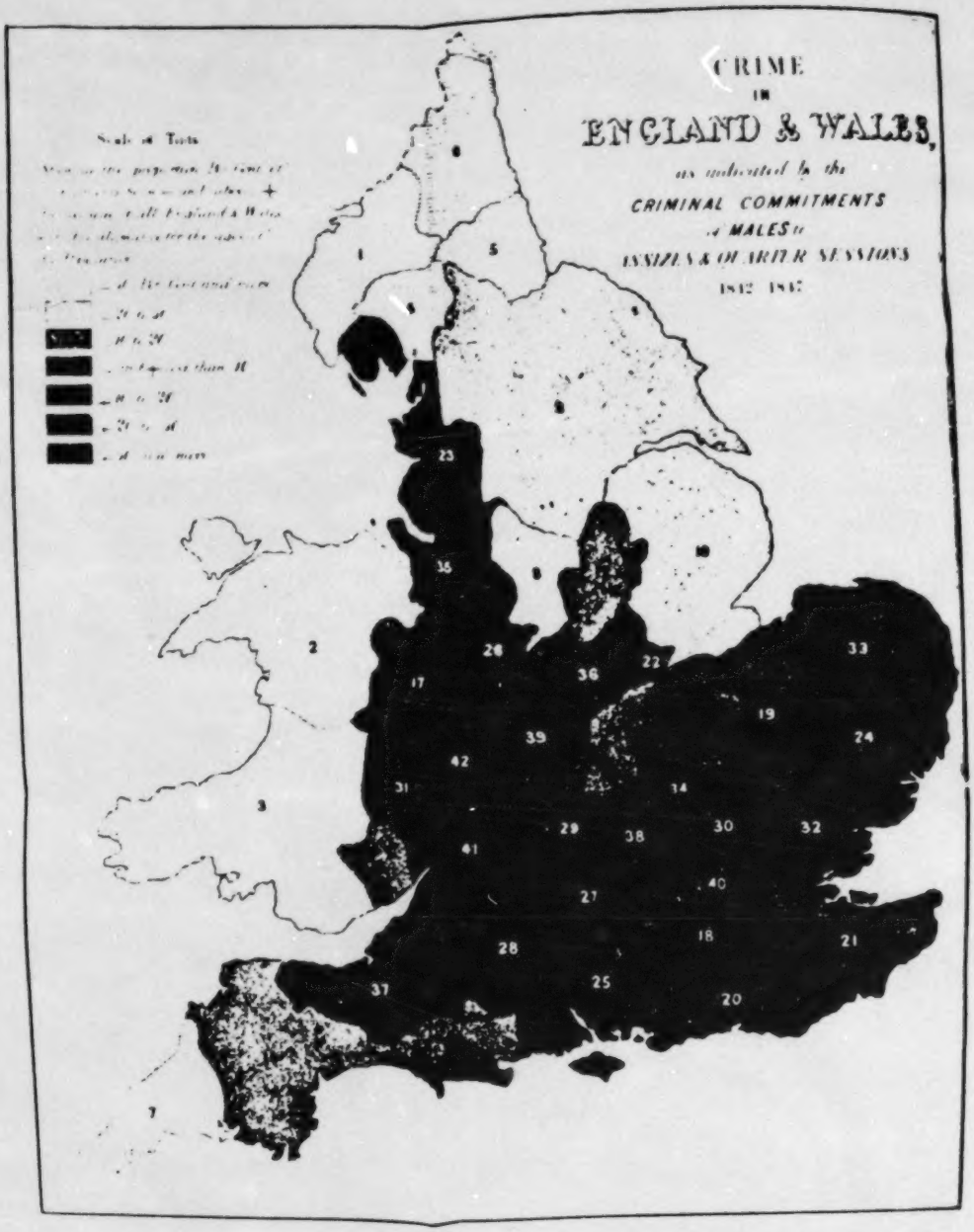

CRIME
IN
ENGLAND & WALES,
as indicated by the
CRIMINAL COMMITMENTS
of MALES
at ASSIZES & QUARTER SESSIONS
1842 1847
Scale of Tints
1
2
3
5
7
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
35
36
37
38
39
40
41
42

Figure 4

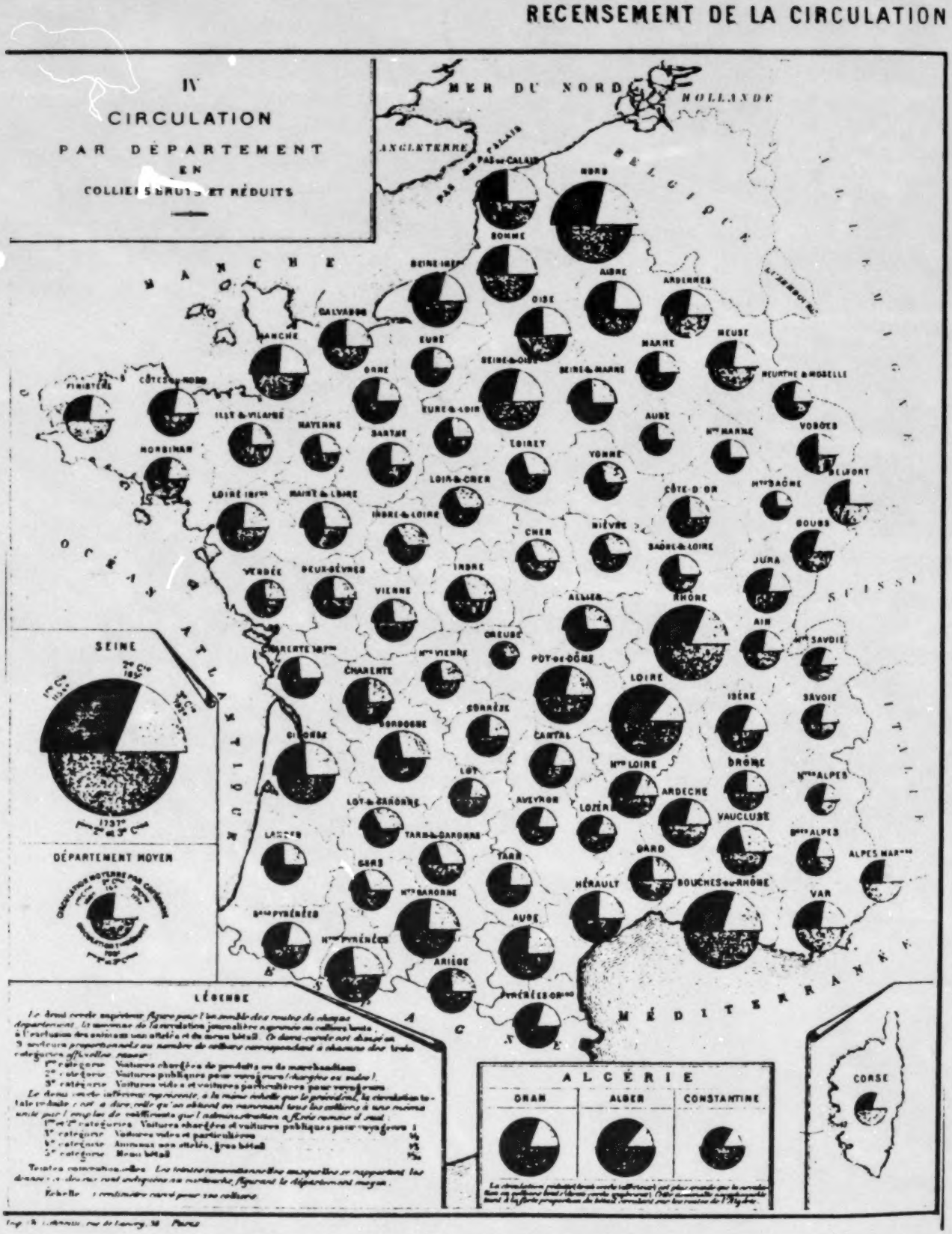

RECENSEMENT DE LA CIRCULATION
IV
CIRCULATION
PAR DÉPARTEMENT
EN
COLLIERS BRUTS ET RÉDUITS
MER DU NORD
HOLLANDE
ANGLETERRE
BELGIQUE
SUISSE
OCÉAN ATLANTIQUE
MÉDITERRANÉE
SEINE
DÉPARTEMENT MOYEN
CORSE
ALGÉRIE
ORAN
ALGER
CONSTANTINE
LÉGENDE

Figure 5

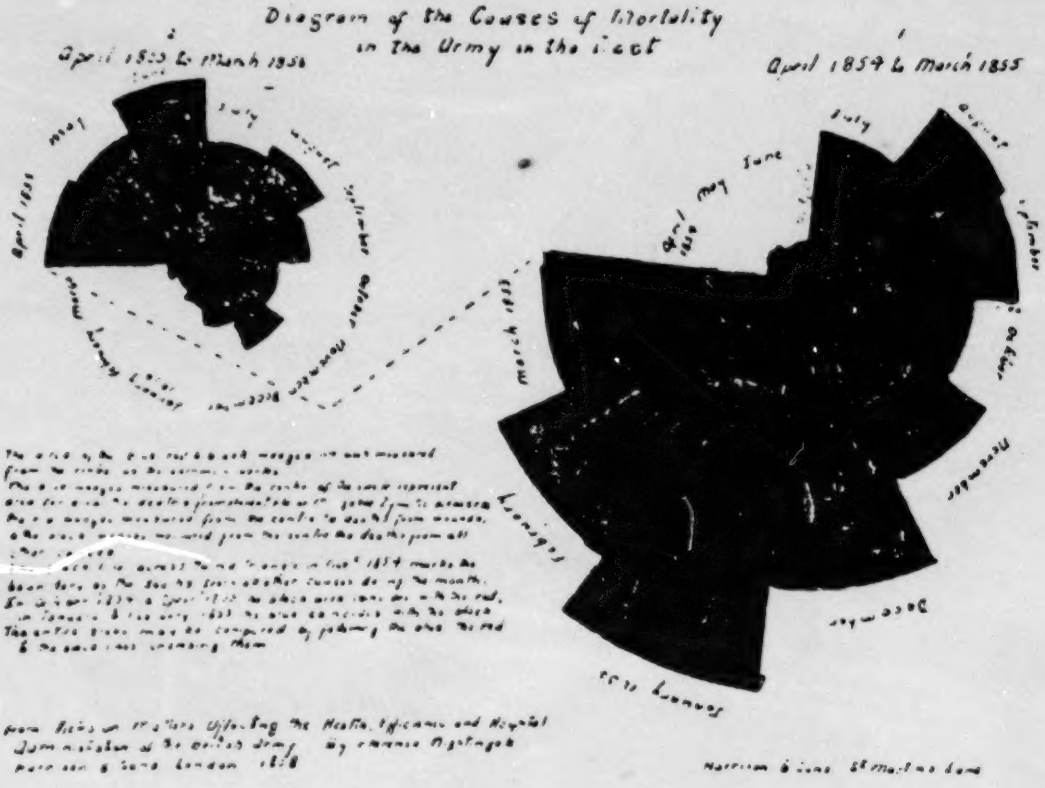


Diagram of the Causes of Mortality
in the Army in the East
April 1854 to March 1855
April 1855 to March 1856
July
August
September
October
November
December
January
February
March
April
May
June

Figure 6

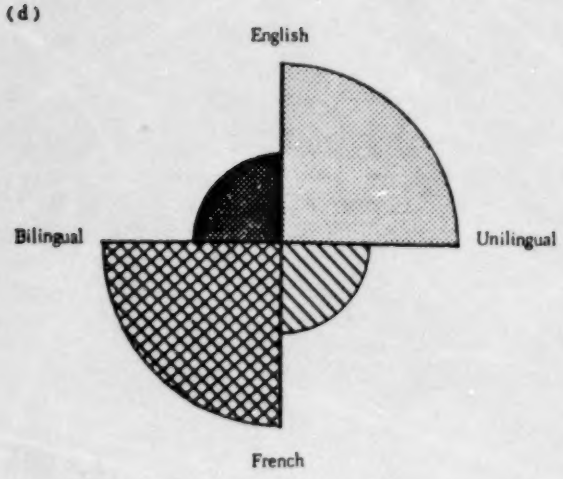
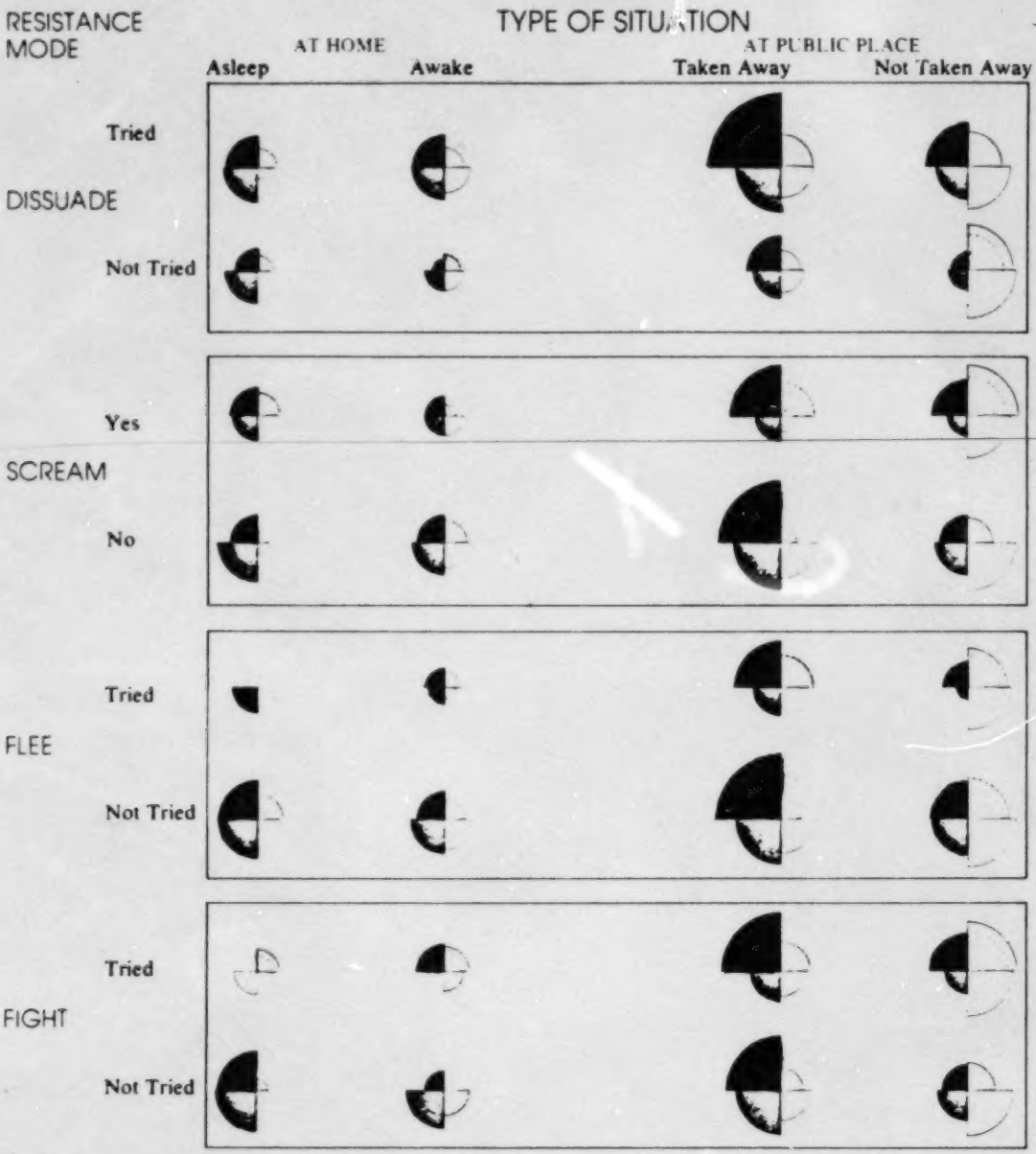


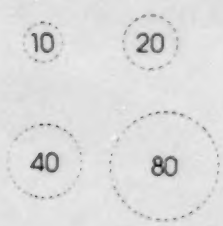
Figure 7

Frequencies of Rapes and Injuries in Various Types of Stranger Rape Situations by Whether or Not a Mode of Resistance Was Used

Quadrantal Tabular Graph



Scale Benchmarks: (Number of cases)



NOTE: Quadrant areas are in proportion to the number of cases. Broken-line circles are proportional to total number of cases in each four-fold figure and, therefore, also show "expected value" of each quadrant. This figure is reproduced here for illustrating the utility of the quadrantal tabular display as a tool for exploratory analysis of cross tabular data. **Caution: The number of cases in most single quadrants is too small for reliability. Modes of resistance are not mutually exclusive. Many informants resisted in more than one way.** The potential utility of this display is that it permits adding together perceptually any sufficiently well-populated combination of elements for whatever comparisons the reader chooses.

Key and Totals:

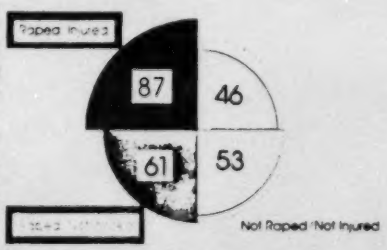


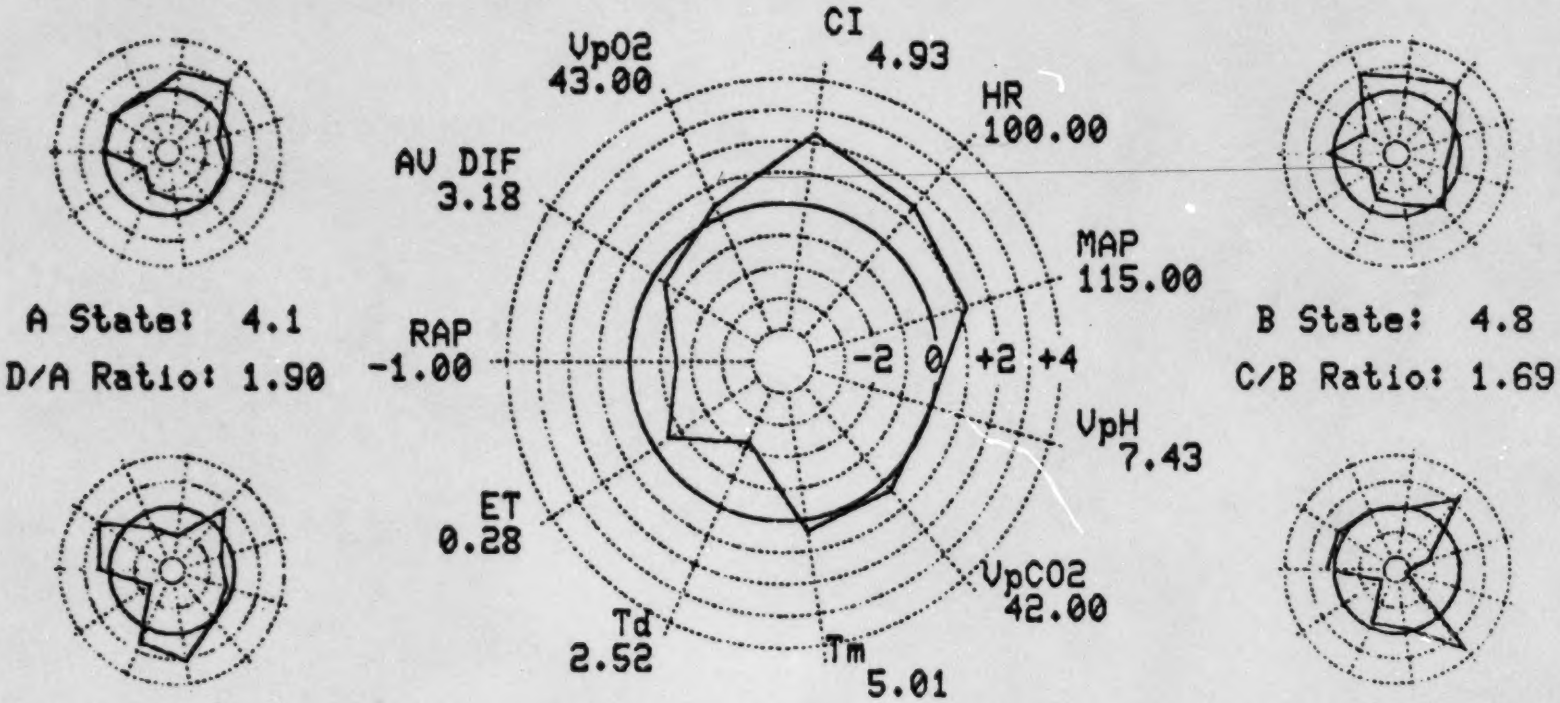
Chart Design: A. D. Biderman and Douglas B. Neal, executed on a Zeta Plotter.

Figure 8

BUFFALO GENERAL HOSPITAL

PATIENT: NUMBER: 848527-01
AGE: 44 YRS SEX:F ADMITTED 1/12/78
ATTENDING PHYSICIAN: SIEGEL

Date 1/16/78 Time 8:25 Intervention No. 6
Height 66.0 in Weight 42.0 kg BSA 1.44 m2
Pat Class R State Dist To R 4.1 Dispersive Vol 298 ml



A State: 4.1
D/A Ratio: 1.90

B State: 4.8
C/B Ratio: 1.69

D State: 7.8

C State: 8.1

Figure 9

Fig. VII

Linien-Diagramme im Kreise.

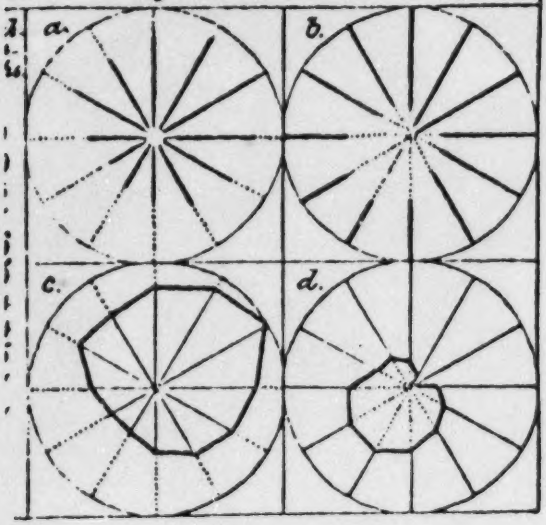


Fig. XII

Flächen-Diagramme im Polygon:

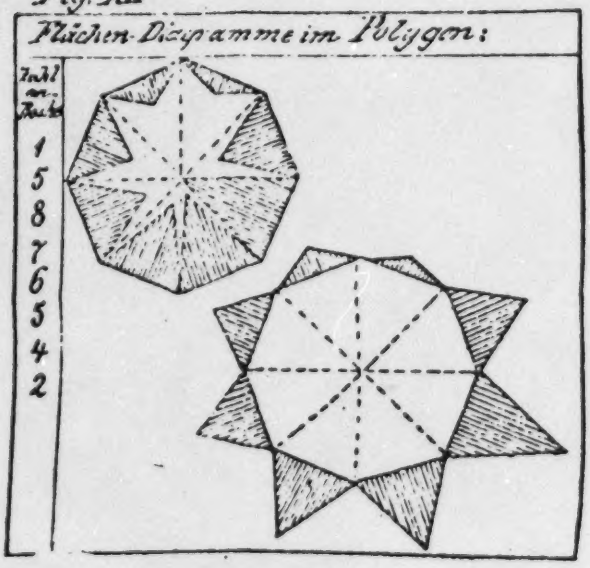
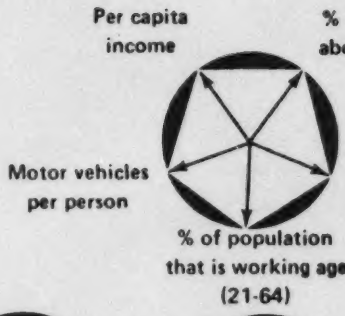


Figure 10

PROFILE OF HEALTH SERVICES USER RESOURCES BY COUNTY
HEALTH SERVICE AREA ONE
(Scaled as fraction of State levels)


Per capita income
% individuals above poverty
Motor vehicles per person
% over 65 not receiving assistance to aged
% of population that is working age (21-64)

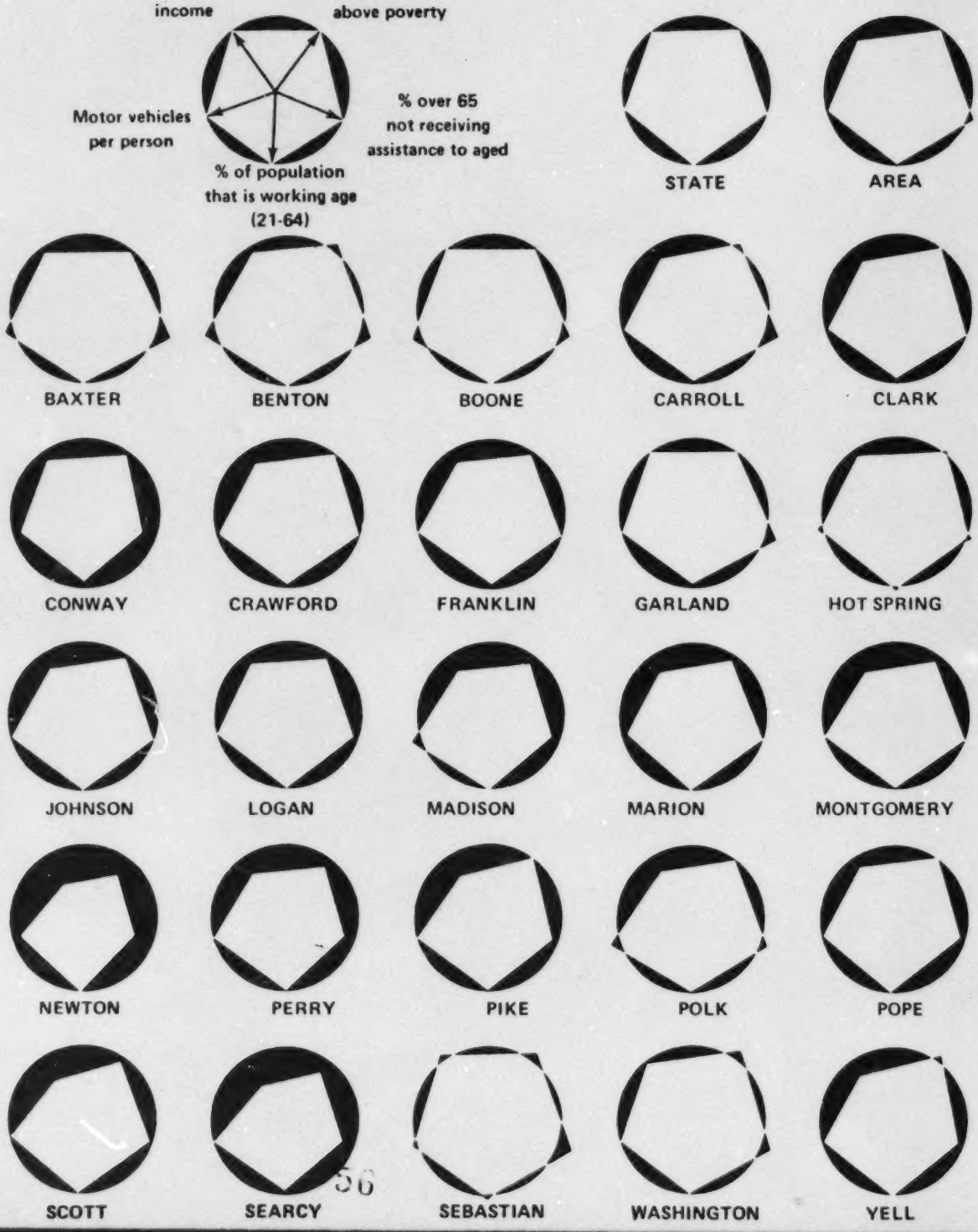


Figure 11

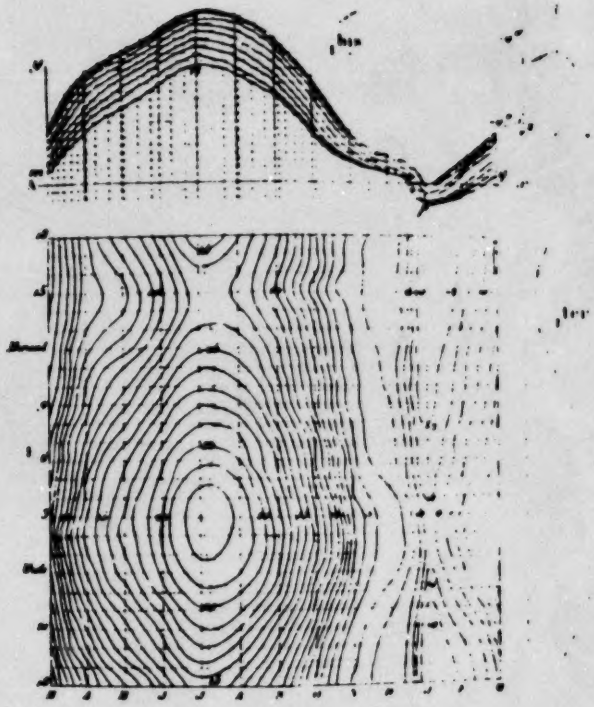


Figure 12

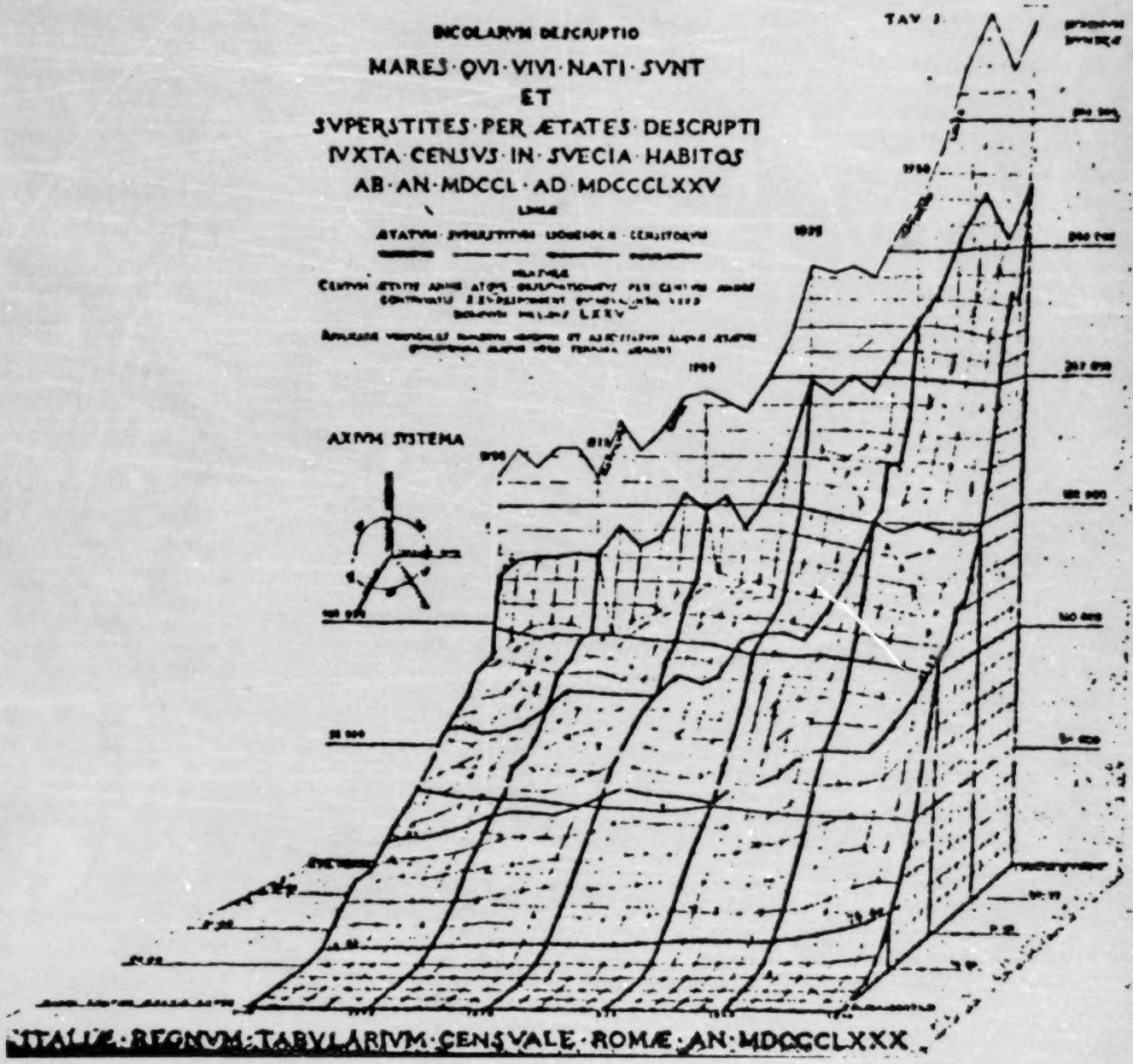

INCOLARVM DESCRIPTIO
MARES·QVI·VIVI·NATI·SVNT
ET
SVPERSTITES·PER ÆTATES·DESCRIPTI
IVXTA·CENSVS·IN·SVECIA·HABITOS
AB·AN·MDCCL·AD·MDCCCLXXV
AXIVM SYSTEMA
ITALIÆ·REGNVM·TABVLARIVM·CENSVALE·ROMÆ·AN·MDCCCLXXX

Figure 13

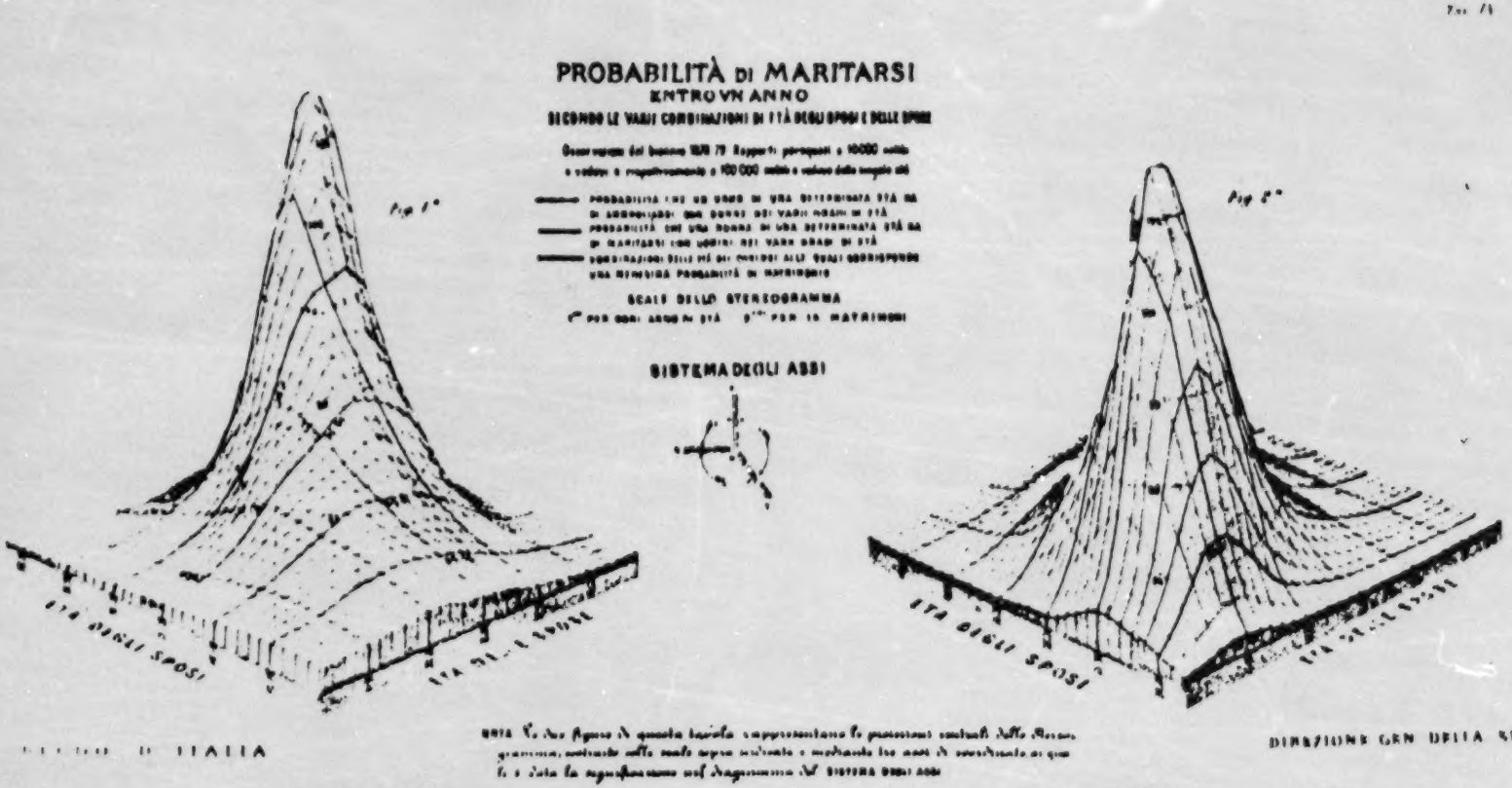

PROBABILITÀ DI MARITARSI
ENTRO UN ANNO
SCALE DELLO STEREOGRAMMA
SISTEMA DEGLI ASSI
ETÀ DEGLI SPOSI
ETÀ DEGLI SPOSI
DIREZIONE GEN DELLA STATISTICA

Figure 14

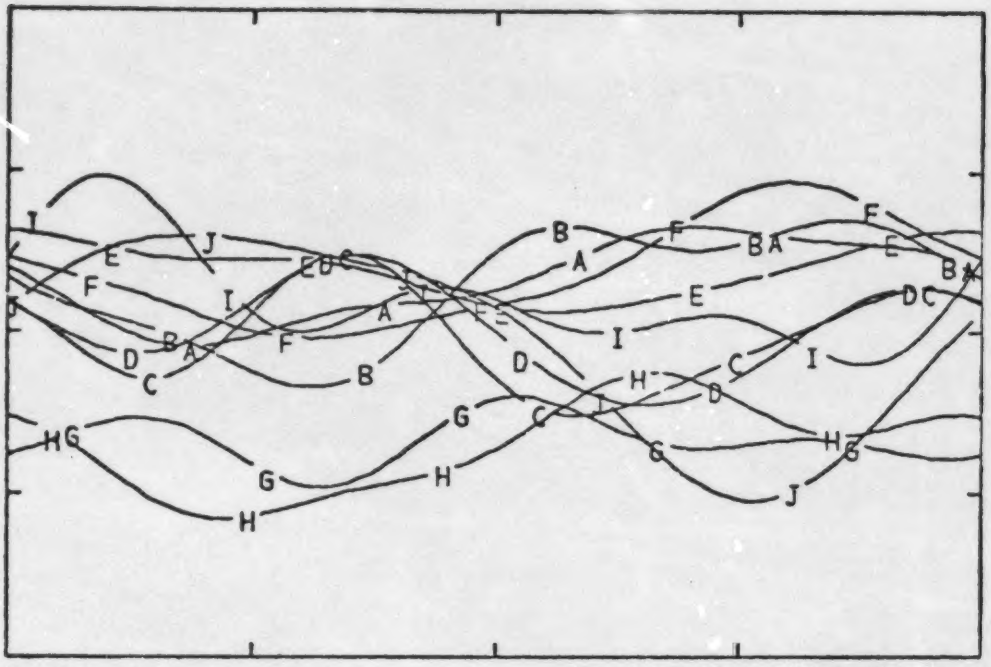
Stem and Leaf Display of State Effects

IOWA, SOUTH DAKOTA, CALIFORNIA	1.0	IA SD CA
OREGON	.6	OR
VERMONT, NEW HAMPSHIRE	.2	VT NH
OHIO, PENNSYLVANIA	-.2	OH PA
	-.6	
	-1.0	
	-1.4	
	-1.8	
	-2.2	
ALABAMA	-2.6	AL
MISSISSIPPI	-3.0	MS

Figure 15

	Life Expectancy	Income	% HS Graduates	Non-Homicide Rate	Literacy Rate	Population	Days Below Freezing
2.8						AL MS	
2.4							
2.0	IA	CA		NH			
1.6				MS			
1.2		OH		SD AL		PA OH	MS
0.8	VT	PA	CA				NH
0.4	OR	OR	OR NH	IA	IA SD VT	CA	AL VT
0.0	SD CA	IA	IA VT	OR VT PA	OR NH OH AL MS	OH	PA OH SD
-0.4		AL	OH AL MS	OH			
-0.8	OH	NH	PA		PA	IA OR	IA
-1.2	NH					NH	
-1.6		SD	SD			VT	
-2.0		VT		CA		SD	
-2.4	AL	MS			CA		OR
-2.8	PA						
-3.2							
-3.6							CA
-4.0							
-4.4	MS						

Figure 16



Function Plot Prepared by David Andrews

Legend

A New Hampshire
B Vermont
C Pennsylvania
D Ohio
E Iowa
F South Dakota
G Alabama
H Mississippi
I Oregon
J California

Figure 17

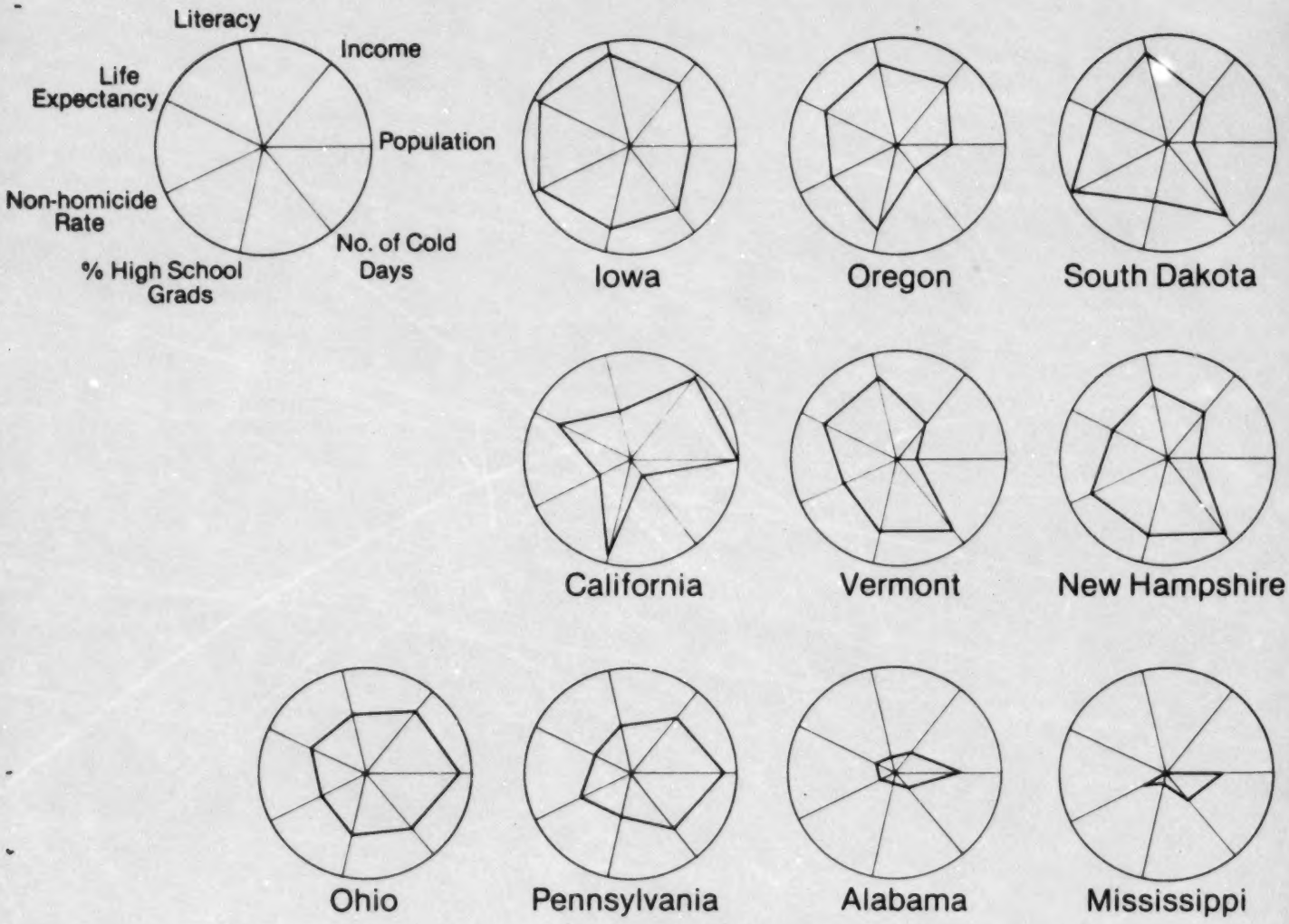

Literacy
Income
Life Expectancy
Population
Non-homicide Rate
No. of Cold Days
% High School Grads
Iowa
Oregon
South Dakota
California
Vermont
New Hampshire
Ohio
Pennsylvania
Alabama
Mississippi

Figure 18

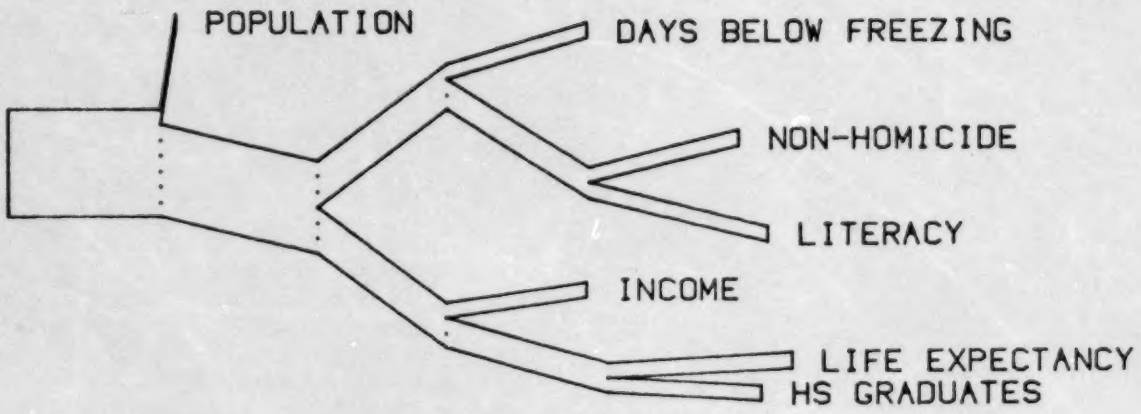
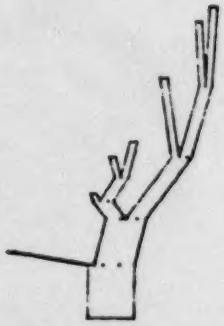

POPULATION
DAYS BELOW FREEZING
NON-HOMICIDE
LITERACY
INCOME
LIFE EXPECTANCY
HS GRADUATES

Figure 19

IOWA



CALIFORNIA

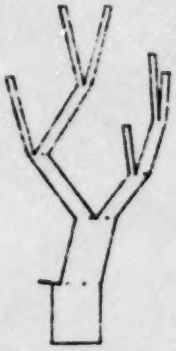
OHIO

MISSISSIPPI

OREGON

VERMONT

PENNSYLVANIA



SOUTH DAKOTA

NEW HAMPSHIRE

ALABAMA

Figure 20

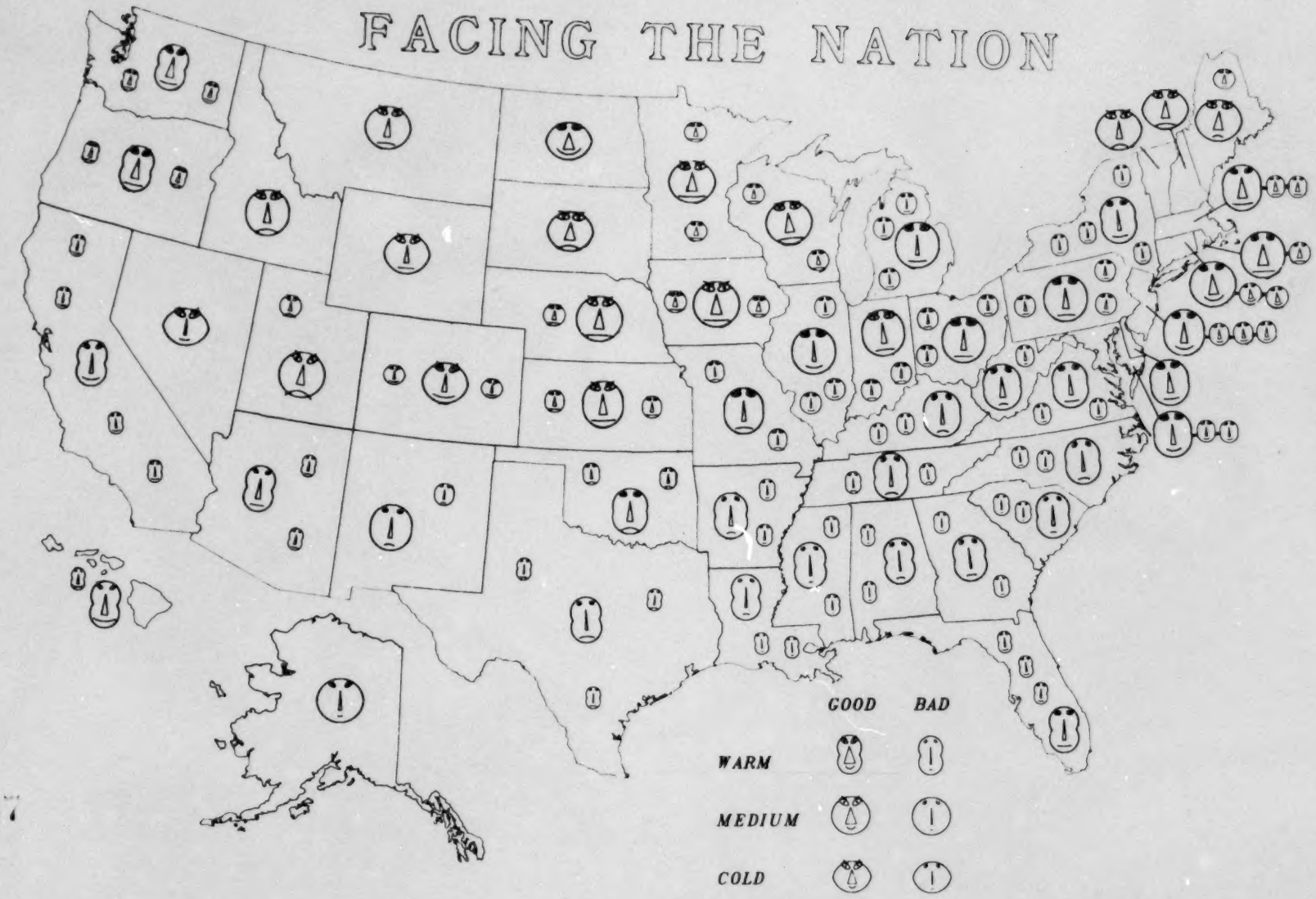


Figure 21

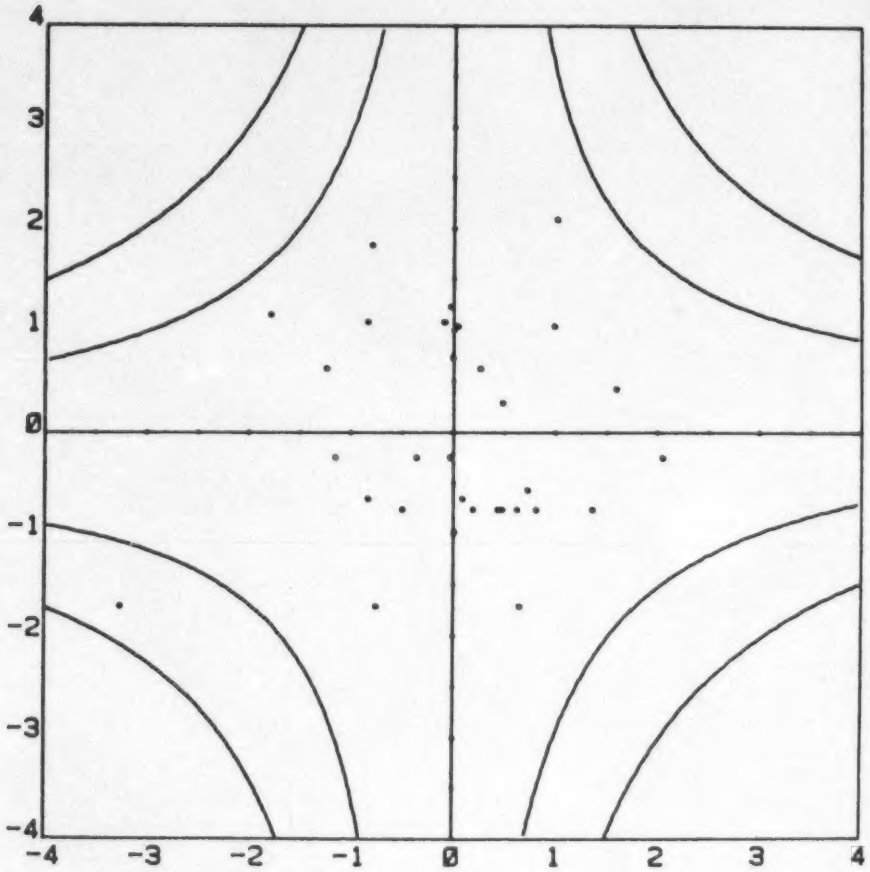


Figure 22

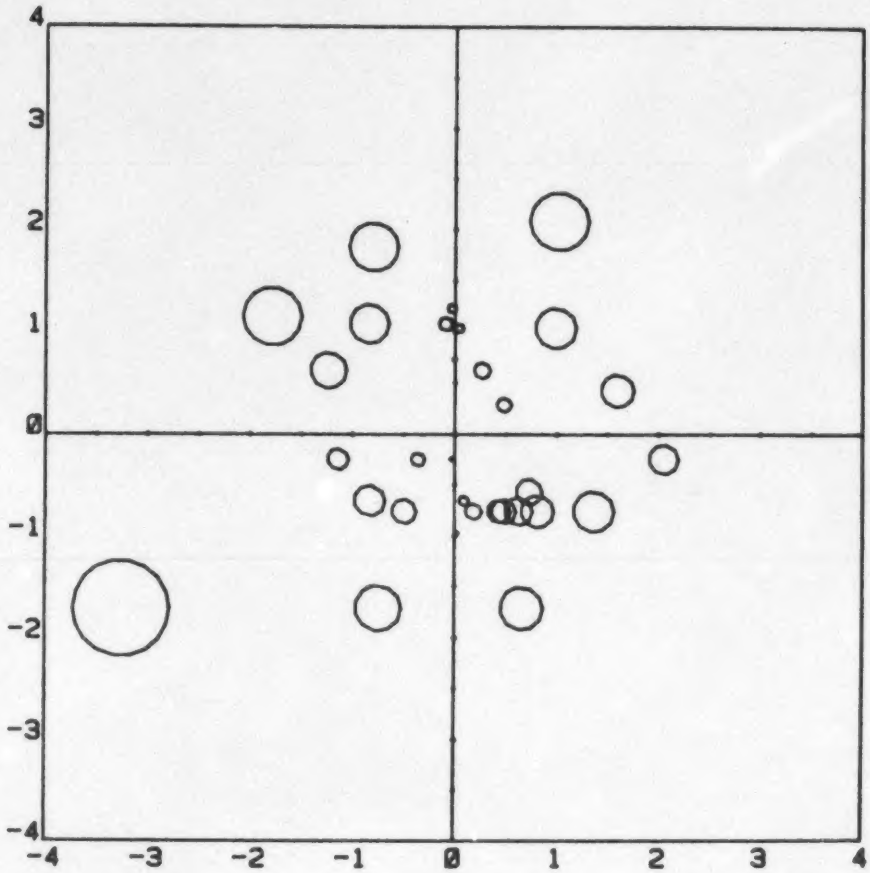


Figure 23

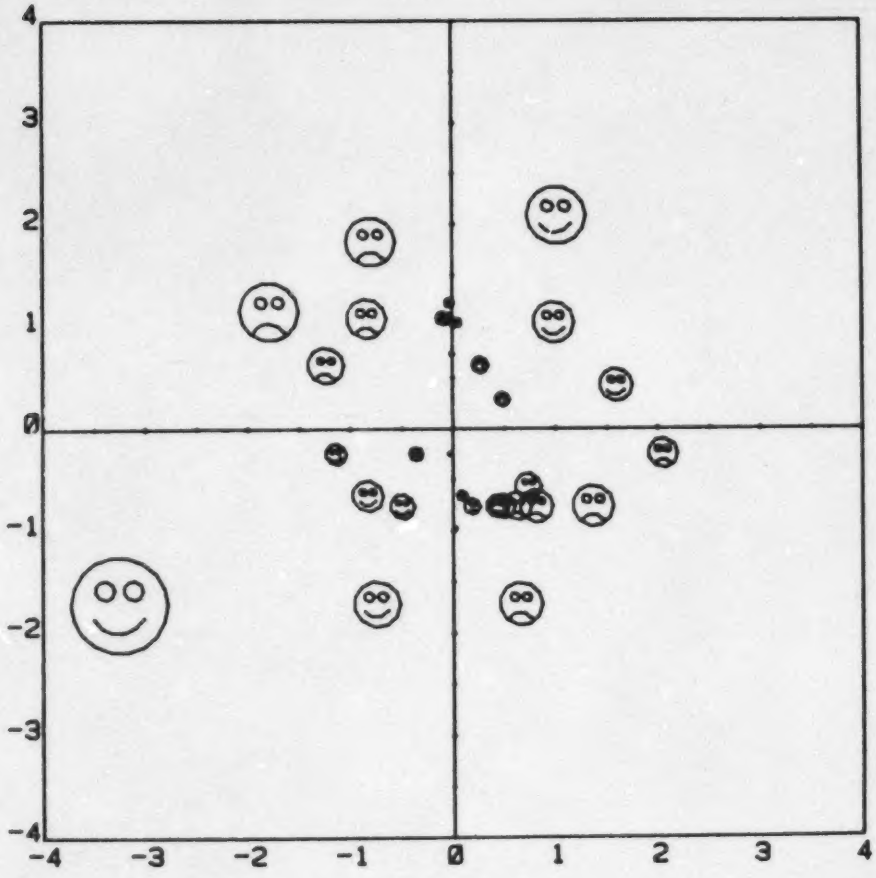


Figure 24

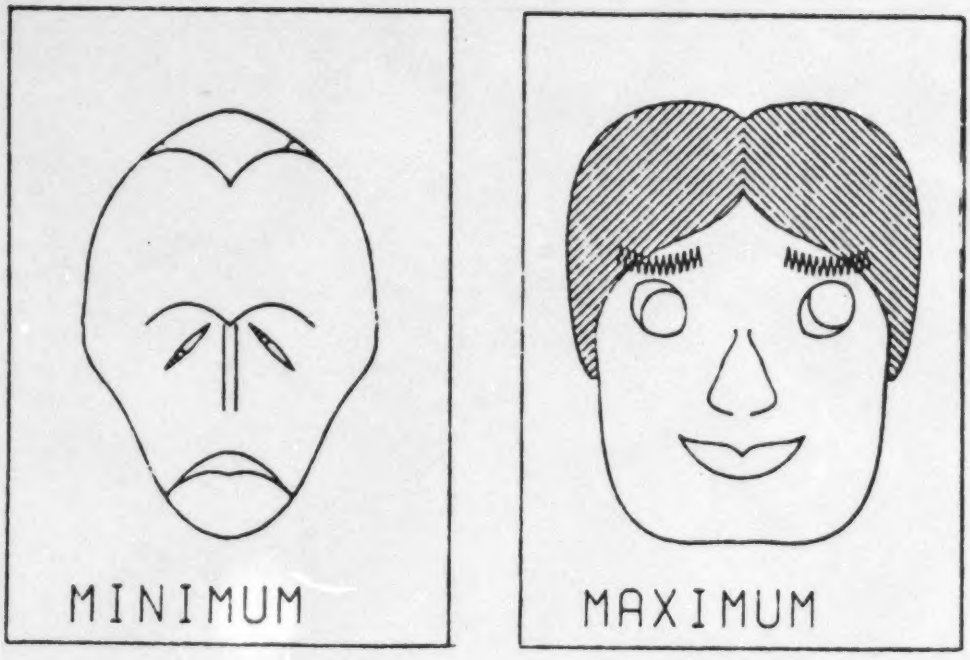


Figure 25

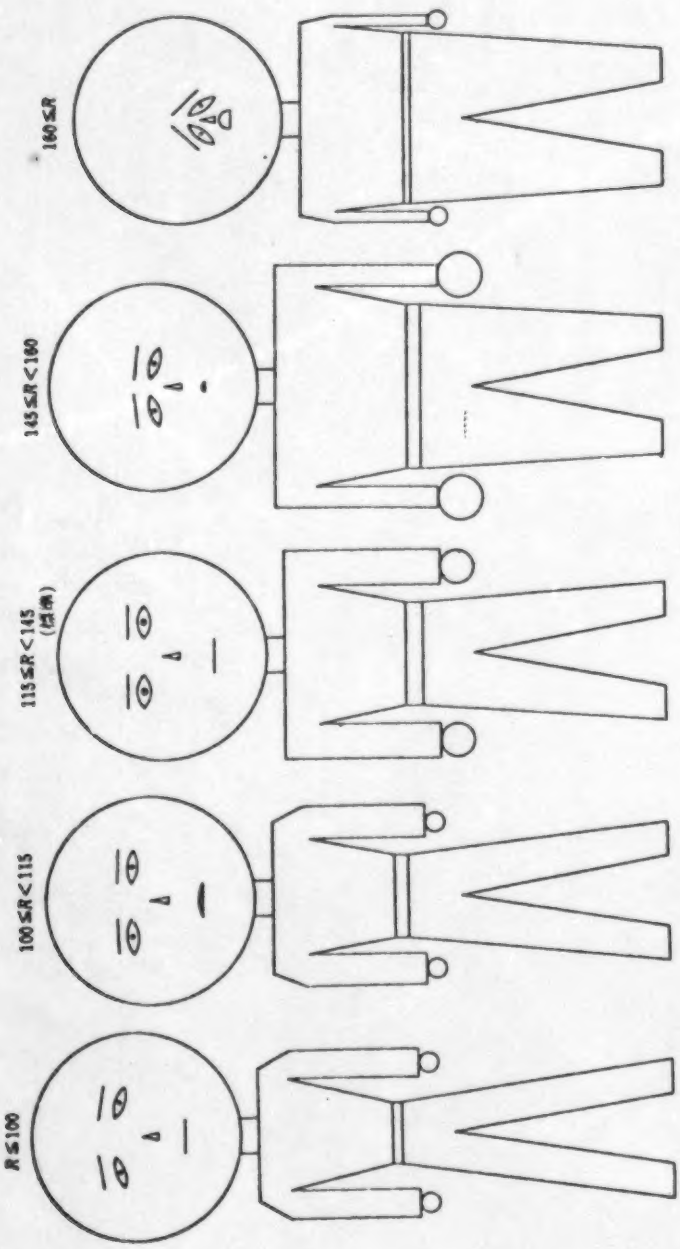


図2 五つのクラスの体形グラフ

Wainer's Wabbits

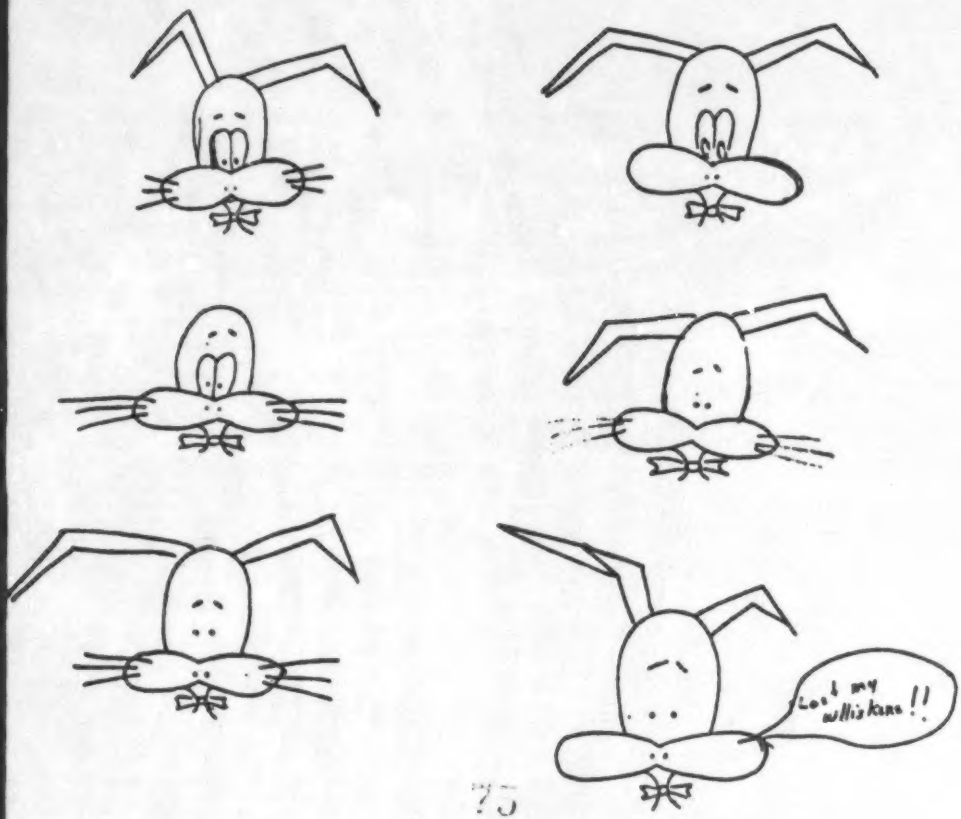


Figure 26

TABLE 1

Excerpted from "Worst American State: Revisited" Data

State	Population (1000's)	Average Per Capita Income ($)	Illiteracy Rate (% Pop.)	Life Expectancy	Homicide Rate (1000)[a]	Percent High School Graduates	Average # Days Year Below Freezing
Alabama	3,615	3,624	2.1	69.05	15.1	41.3	20
California	21,198	5,114	1.1	71.71	10.3	62.6	20
Iowa	2,861	4,628	.5	72.56	2.3	59.0	140
Mississippi	2,341	3,098	2.4	68.09	12.5	41.0	50
New Hampshire	812	4,281	.7	71.23	3.3	57.6	174
Ohio	10,735	4,561	.8	70.82	7.4	53.2	124
Oregon	2,284	4,660	.6	72.13	4.2	60.0	44
Pennsylvania	11,860	4,449	1.0	70.43	6.1	50.2	126
South Dakota	681	4,167	.5	72.08	1.7	53.3	172
Vermont	472	3,907	.6	71.64	5.5	57.1	168

[a]"Homicide Rate/1000" is actually combined murders and non-negligent manslaughter per 1000 population.

TABLE 2

TABLE 1 Rounded with Rows and Columns Reordered and Closed Up

State	Life Expec- tancy	In- come (100's)	% HS Grad- uates	Homi- cide Rate	Illit- eracy Rate	Population (100,000's)	Days Below Freezing
Iowa	73	46	59	2	.5	29	140
Oregon	72	47	60	4	.6	23	44
South Dakota	72	42	53	2	.5	7	172
California	72	51	63	10	1.1	212	20
Vermont	72	39	57	6	.6	5	168
New Hampshire	71	43	58	3	.7	8	174
Ohio	71	46	53	7	.8	107	124
Pennsylvania	70	44	50	6	1.0	119	126
Alabama	69	36	41	15	2.1	36	20
Mississippi	68	31	41	12	2.4	31	50
Median	72	44	55	6	.8	26	125

TABLE 3

The Same Data as in TABLE 2 after Transformations

State	Life Expec- tancy	In- come	% HS Grad- uates	Homi- cide Rate	Lit- eracy Rate	Popu- lation	Days Below Freezing
Iowa	73	68	59	30	141	14.9	140
Oregon	72	68	60	27	129	14.6	44
South Dakota	72	65	53	32	141	13.4	172
California	72	72	63	23	95	16.9	20
Vermont	72	63	57	26	129	13.1	168
New Hampshire	71	65	58	29	120	13.6	174
Ohio	71	68	53	24	112	16.2	124
Pennsylvania	70	67	50	25	100	16.3	126
Alabama	69	60	41	21	69	15.1	20
Mississippi	68	56	41	22	65	14.7	50
Median	71.5	66	55	25.5	116	14.8	125

TABLE 4

The Same Data as in TABLE 3 Column-Centered by Subtracting Out Medians

State	Life Expec- tancy	In- come	% HS Grad- uates	Non- Homicide Rate	Lit- eracy Rate	Popu- lation	Days Below Freezing
Iowa	1.5	2	4	4.5	25	.1	15
Oregon	.5	2	5	1.5	13	- .2	- 81
South Dakota	.5	- 1	- 2	6.5	25	-1.4	47
California	.5	6	8	-2.5	-21	2.1	-105
Vermont	.5	- 3	2	.5	13	-1.7	43
New Hampshire	- .5	- 1	3	3.5	4	-1.2	49
Ohio	- .5	2	- 2	-1.5	- 4	1.4	- 1
Pennsylvania	-1.5	1	- 5	- .5	-16	1.5	1
Alabama	-2.5	- 6	-14	-4.5	-47	.3	-105
Mississippi	-3.5	-10	-14	-3.5	-51	- .1	- 75
Median	71.5	66	55	25.5	116	14.8	125
MAD	.5	2	4.5	3	18	1.3	45

TABLE 5

The Same Data as in TABLE 4 Column-Centered and Column-Standardized by the MAD

State	Life Expec-tancy	In-come	% HS Grad-uates	Non-Homicide Rate	Lit-eracy Rate	Popu-lation	Days Below Freezing	Row Median
Iowa	3	1.0	.9	1.5	1.4	.1	.3	1.0
Oregon	1	1.0	1.1	.5	.7	- .2	-1.8	.7
South Dakota	1	- .5	- .4	2.2	1.4	-1.1	1.0	1.0
California	1	3.0	1.8	- .8	-1.2	1.6	-2.3	1.0
Vermont	1	-1.5	.4	.2	.7	-1.3	1.0	.4
New Hampshire	-1	- .5	.7	1.2	.2	-.9	1.1	.2
Ohio	-1	1.0	- .4	- .5	-.2	1.1	0	- .2
Pennsylvania	-3	.5	-1.1	- .2	-.9	1.2	0	- .2
Alabama	-5	-3.0	-3.1	-1.5	-2.6	.2	-2.3	-2.6
Mississippi	-7	-5.0	-3.1	-1.2	-2.8	- .1	-1.7	-2.8
Median	71.5	66	55	25.5	116	14.8	125	
MAD	.5	2	4.5	3	18	1.3	45	

TABLE 6

The Same Data as in TABLE 5, Column-Scaled and Doubly Centered by Medians--A Scaled Residual Matrix

State	Life Expec-tancy	In-come	% HS Grad-uates	Non-Homicide Rate	Literacy Rate	Popu-lation	Days Below Freezing
Iowa	2.0	0	- .1	.5	.4	-.9	- .7
Oregon	.3	.3	.4	- .2	0	-.9	-2.5
South Dakota	0	-1.5	-1.4	1.2	.4	-2.1	0
California	0	2.0	.8	-1.8	-2.2	.6	-3.3
Vermont	.6	-1.9	0	- .2	.3	-1.7	.6
New Hampshire	-1.2	-.7	.5	1.0	0	-1.1	.9
Ohio	- .8	1.2	- .2	- .3	0	1.3	.2
Pennsylvania	-2.8	.7	- .9	0	- .7	1.4	.2
Alabama	-2.4	-.4	- .5	1.1	0	2.8	.3
Mississippi	-4.2	-2.2	- .3	1.6	0	2.7	1.1